DYNAMICS OF POINT AND LINE DEFECTS IN
SINGLE SEMICONDUCTOR CRYSTALS GROWN FROM THE MELT

by

DIMITRIOS MAROUDAS

Diploma in Chemical Engineering
National Technical University of Athens (1987)

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

September, 1992

© Massachusetts Institute of Technology 1992. All rights reserved.

Signature of Author

Department of Chemical Engineering
July 14, 1992

Certified by

Robert A. Brown
Professor, Department of Chemical Engineering
Thesis Supervisor

Accepted by

Robert E. Cohen
Professor, Department of Chemical Engineering
Chemical Engineering Graduate Officer

NOV 02 1992
LIBRARIES
ARCHIVES
DYNAMICS OF POINT AND LINE DEFECTS IN SINGLE SEMICONDUCTOR CRYSTALS GROWN FROM THE MELT

by

DIMITRIOS MAROUDAS

Submitted to the Department of Chemical Engineering on July 14, 1992, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering

ABSTRACT

Single crystals of elementary and compound semiconductors are the material substrates for electronic and optoelectronic device fabrication technologies. The quality of these crystals is determined by defects in the crystalline lattice, associated with crystal growth and device processing. Line defects or dislocations are lattice imperfections with the most catastrophic effects on the performance of semiconductor devices. Point defects either can enhance or degrade device performance. For example, some impurity atoms cause lattice hardening and significantly improve the mechanical properties of the crystals, while supersaturated vacancies and self-interstitials form clusters that provide sources for dislocation formation. The objective of this thesis is the establishment of a theoretical and computational framework for the quantitative modeling of defect dynamics on the atomic, microscopic, and continuum scales. Quantitative theoretical predictions provide the basis for the analysis of the effects of materials processing conditions on the type and density of crystallographic defects, so that strategies to improve the material quality can be developed.

Dislocation dynamics in a growing crystal is analyzed using the model of Alexander and Haasen (1968) for plastic deformation in diamond/zincblende structure semiconductors. The thermal stress fields used in the analysis are indicative of the liquid encapsulated Czochralski (LEC) growth of InP and GaAs and of the Czochralski (CZ) growth of silicon. For typical crystal growth conditions, the rapid rate of dislocation multiplication compared to the rate of motion of the crystal through the thermal stress field leads to asymptotic expressions for the bulk dislocation density as a function of the variation of the stress field in the direction of crystal growth. The asymptotic results agree quantitatively with computations and experimental results for dislocation densities in LEC grown InP and give simple criteria for reducing dislocation densities in as-grown compound semiconductor crystals.

The interaction between impurities and dislocations is studied for the case of interstitial oxygen in CZ grown silicon. A systematic theory is developed for the analysis of oxygen migration near a 60° dislocation and for the resulting retardation of dislocation motion. Quantitative predictions are based on the solution of the mesoscale equation for oxygen transport in the elastic stress field created by the dislocation. The oxygen concentration field is computed numerically using the Galerkin finite-element method. The link between the microscopic dynamics of interstitial oxygen within the diamond lattice and
macroscopic transport is established by a constitutive model for the dependence of the drift velocity and diffusivity of oxygen on the elastic interaction and temperature. Drag on dislocation motion is modeled using the oxygen distribution around the dislocation core, the elastic interaction forces and the energy needed to dissolve oxygen clusters that are formed within the core. The model predicts dislocation locking due to oxygen. The predictions of the model agree quantitatively with experimental data for dislocation mobility in silicon grown by the Czochralski and floating zone methods (Imai and Sumino, 1983). Constitutive models also have been developed that describe the effects of oxygen on the internal stresses in the crystal and on the multiplication of dislocations due to the cross-slip mechanism. Combining the results of the dislocation drag model with the constitutive models for lattice hardening and dislocation multiplication reproduces experimentally measured stress-strain curves (Yonenaga et al., 1984) and explains the initial stage of deformation of oxygen-hardened silicon.

Thermodynamic, structural and transport properties of intrinsic point defects, i.e. vacancies and self-interstitials, in crystalline silicon are predicted using an integrated approach based on atomistic simulations. Crystalline silicon is modeled by the Stillinger-Weber interatomic potential (Stillinger and Weber, 1985). Lattice statics, lattice dynamics, and Monte Carlo (MC) simulations are combined to calculate crucial point defect parameters needed in mesoscale models of defect dynamics. Thermal equilibrium concentrations of point defects are calculated from Gibbs free energy computations based on a combination of the quasi-harmonic approximation and the analysis of the first few cumulants of the enthalpy distribution function computed from isothermal-isobaric MC simulations. The temperature dependence of formation enthalpies and volumes of point defects are also computed using the same approach. The zero-temperature structure and energetics of point defects also are studied using a combination of the simulated annealing and the steepest descent methods for energy minimization. Important point defect parameters, such as formation energies and activation energies of migration, are calculated. Diffusion coefficients of vacancies and self-interstitials are computed based on jump rate theory and the simulation results for enthalpies and Gibbs free energies of migration. Migration energies and thermal equilibrium concentrations of vacancies are in reasonable agreement with experimental data. Formation and migration properties of self-interstitials are in good agreement with experimental data. The calculations reproduce experimentally measured interstitial diffusivities and explain interstitial contributions to self-diffusion and substitutional dopant diffusion.

The robust computational framework developed here for the lattice-scale study of point defects in silicon provides the necessary background for a systematic investigation of mechanisms for the formation of larger scale defects such as faulted loops and dislocations. The nucleation and growth of clusters formed by intrinsic point defects under supersaturation is the most important mechanism for defect formation.

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

Having spent five difficult but memorable years at MIT I feel very happy that I am in a position to acknowledge the contribution of my teachers, colleagues, and many friends to the completion of this thesis. I will never forget this period of my life, which was full of knowledge and experience gained by immeasurable effort. The product of most of this effort is presented in this thesis.

It has been a privilege for me to conduct this research under the guidance of Professor Robert A. Brown. Bob’s infinite energy, his enthusiasm about science and quantitative research, and his intelligence has been and will continue to be an inspiration. Bob was not only my thesis advisor but also an excellent teacher of transport phenomena, numerical methods and nonlinear analysis, which were my original fields of interest that made come to the United States for graduate studies. Bob was a general mentor for me during the past five years. His consistency on high standards for theoretical and computational research, teaching, writing and presenting scientific work has influenced my point of view as a member of the scientific and academic community significantly. The last but definitely not the least acknowledgement of Bob’s contribution and assistance is for providing the necessary funding that allowed me to use state-of-the-art computer facilities, which made most of this research possible. Many calculations of point defect properties in silicon were performed on the CRAY XMP maintained by the MIT Supercomputer Facility. Funding provided by the Materials Processing Center at MIT in terms of a fellowship also is greatly acknowledged.

The interdisciplinary nature of this study gave me the opportunity to interact with many faculty members at MIT. I am obliged to the members of my thesis committee: Professors Robert C. Armstrong and Klavs F. Jensen of the Department of Chemical Engineering, Professor Robert W. Balluffi of the Department of Materials Science and Engineering and Professor Sidney Yip of the Department of Nuclear Engineering. All of them were very friendly, willing to discuss research topics, and very helpful especially during my last academic year at MIT. It is not at all coincidental that from the numerous presentations I had to give during the past five academic years my thesis committee meetings were the
most enjoyable. It has been a privilege for me to have Professor Balluffi as my first teacher of "defects in crystals". Several discussions with Professor Yip on materials physics and computational statistical mechanics were always enlightening and fruitful. My common scientific interests with Professor Yip have led to a personal and professional friendship, which I hope will be longlasting. I also want to thank Professor John D. Joannopoulos of the Department of Physics for useful discussions about defects in silicon and for making available some interesting results for oxygen in silicon prior to publication. Finally, I want to thank Professor Howard Brenner of the Department of Chemical Engineering for being an excellent teacher of applied mathematics and continuum mechanics.

I consider myself very lucky that the many hours I spent at the office, Room 66-256, were made pleasant by the company and friendship of my colleagues including Angelo Kyr- lidis, Costas Tsiveriotis, Tom Kinney, Howard Covert, Reza Mehrabi, Alparslan Oztekin, Chuck Krueger, and Talid Sinno. Fortunately enough we managed to coexist in places other than 66-256. I am very confident that all of them will enjoy professional success and I wish them the best in their personal lives. Angelo and Costas were my earliest friends at MIT. Angelo has been a friend for a long time starting from our undergraduate years in Athens. We have shared many good and bad times being classmates, roommates, and officemates for the past five years and having many common scientific and other interests including the world of the movies. Angelo, I wish you the best of luck in your professional endeavors. Costas has always been a true friend. His willingness to help me when I arrived in Boston one year after him and the many discussions we had at the office will never be forgotten. Discussions with Tommie looking at the world map can not be forgotten either. The last two months that I spent writing this thesis were definitely made easier by the presence of Reza, Howard, Alp, Talid, and especially Chuck (the "Solarian" King) at the office. Talid, I wish you the best of luck in the continuation of exploring the world of defects in silicon. Also, I want to thank some of the "older" members of our group including Mark Bennett, Sanjay Patnaik, Do Hyun Kim, Narayanan Ramprasad, and John Lee for sharing some of their experience during my early days in the group. Finally, I want to say a big thanks to Arline Benford for her kindness and her constant willingness to help me with various administrative procedures. Arline, I will really miss you.
It is very difficult to thank all the friends that made life happier during the past five years. Many special thanks go to Costas Douzinias, Nadia Dimou, Anna Protopappas, Stavros Karatzoulas, and Ioanna Petrounia for all the good times we spent together here in Boston. From the endless list of friends back in Greece I would like to greatly acknowledge Thanasis Binis, Takis Meggos, Niki Niktari, and Dionisis Mavrotas for keeping in touch and for making my summer vacations really enjoyable. From the many friends here at MIT I would like to specially acknowledge Dan Coronell, Todd Salamon, Leo Lue, and Joy Mendoza for making my participation in several basketball games during the spring and the summer of 1991 a learning experience. Finally, I want to say a big thanks to Tony Prentakis, the best friend I made here in Boston, for his support, his intelligence and sense of humor, and the many good times we shared at Harvard Square and Back Bay.

It is really difficult to adequately thank my parents, Apostolos and Despina Maroudas, for their great love and the strong support that they always offered me in every difficulty I had to face. I feel very proud of my parents for the way they brought me up and for their effort for me to obtain a good education, as a result of their belief in the importance and the power of knowledge. I am sure that the completion of this thesis makes them happy. Many thanks also go to my sister, Eleni, and my brother-in-law, Vangelis Platsis, for their support, the many phone calls, and their effort to make my vacation at home more pleasant. Following the growing up of my lovely nephews, Stergios and Akis, has been one of the best experiences of the past few years. Although these two amazing guys are very young to understand what I mean, I thank them for making their parents and grandparents happy.

Finally and most importantly, I would like to thank Mary Prentakis, the person who made the most significant impact in my life during the past few years. I really feel so very lucky that I met you Mary. Thank you for your love, your patience, your strength, and your constant and infinite support in so many ways. The most precious gifts that I received during these years were the pleasure of your company and your mature point of view about life and happiness. I want to believe that the completion of this thesis marks the end of the hardest period and the beginning of the happier times of our lives together. Mary, this thesis is dedicated to you.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ 11

LIST OF TABLES ........................................................................................................... 22

1. INTRODUCTION ...................................................................................................... 23
   1.1. Point and Line Defects in Single Crystals .................................................... 23
   1.2. Crystal Growth and Defect Generation ....................................................... 25
   1.3. Modeling Philosophy ..................................................................................... 30
   1.4. Thesis Objectives and Outline .................................................................... 34

2. DISLOCATION FORMATION AND DYNAMICS IN SEMICONDUCTOR
   CRYSTALS ................................................................................................................. 36
   2.1. Dislocations and Mechanical Strength of Single Crystals ......................... 36
       2.1.1. Crystallographic Glide and Dislocations ............................................. 36
       2.1.2. Critical Resolved Shear Stress (CRSS) .............................................. 41
       2.1.3. Dislocation Motion and Plastic Flow .................................................. 45
   2.2. Dislocation Dynamics in the Diamond/Zincblende Structure .................. 46
       2.2.1. Mobile Dislocation Configuration ....................................................... 46
       2.2.2. Dislocation Mobility in High-Purity Crystals ...................................... 47
       2.2.3. Dislocation Multiplication .................................................................. 51
   2.3. Effects of Impurities on Dislocation Mobility ............................................. 53
   2.4. Dislocation Generation During Crystal Growth ......................................... 61
       2.4.1. Mechanisms of Dislocation Formation During Growth ..................... 62
       2.4.2. Formation of Dislocations by Point Defect Clusering ....................... 64
       2.4.3. Experimental Evidence for Dislocation Loop Formation in Silicon
               and III-V Compounds ........................................................................ 66
   2.5. Dislocation Modeling in Crystal Growth ..................................................... 68
2.5.1. The CRSS Model for Dislocation Formation ........................................68
2.5.2. Advances in Thermal Stress Modeling for CZ and LEC Growth ..........71
2.5.3. Modeling of Inelastic Effects ................................................................72

3. QUANTITATIVE DISLOCATION MODELING IN CRYSTAL GROWTH ....74

3.1. Dislocation Structure and Dynamics in Diamond/Zincblende
   Structure Crystals ......................................................................................74
3.1.1. Microscopic Aspects ..........................................................................74
3.1.2. Atomistic Simulations of Dislocation Structure and Dynamics ..........78
3.1.3. Macroscopic Model of Dislocation Dynamics .....................................81

3.2. Modeling of Dislocation Dynamics During Semiconductor Crystal Growth ......86
3.2.1. The Model of Alexander and Haasen .................................................88
3.2.2. Parameter Estimates ...........................................................................91
3.2.3. Plastic Deformation Dynamics of a Growing Crystal .......................94
   3.2.3.1. Mathematical Formulation for the Calculation of the
           Thermal Elastoplastic Stress Field ....................................................95
   3.2.3.2. Model of Plastic Deformation Dynamics During Growth ..........96

3.3. Scaling Analysis and Asymptotic Calculation of the Dislocation Density ........100
3.3.1. Scaling of Dislocation Density and Plastic Strain ..............................100
3.3.2. Constant von Mises Stress: \( \sigma(\zeta) = 1 \) ........................................101
3.3.3. Spatially Varying von Mises Stress: \( \sigma = \sigma(\zeta) \) ............................107
   3.3.3.1. Decreasing von Mises Stress: \( d\sigma/d\zeta < 0 \) ..............................108
   3.3.3.2. Increasing von Mises Stress: \( d\sigma/d\zeta > 0 \) ..............................109
3.3.4. Accuracy of Asymptotic Results ........................................................111

3.4. Implications for Czochralski and Liquid Encapsulated Czochralski
   Crystal Growth ......................................................................................117
3.4.1. Growth of InP and GaAs .................................................................117
3.4.2. CZ Growth of Si ..............................................................................122
3.4.3. Effects of Necking During Seeding: Si versus InP ............................124
3.5. Summary and Discussion .................................................................................................................. 125

4. TRANSPORT OF INTERSTITIAL OXYGEN AROUND DISLOCATIONS IN SILICON ................................................................................................................................. 130

4.1. Constitutive Modeling of Interstitial Impurity Diffusion and Drift .............................................. 131
   4.1.1. Jump Rate Theory for Point Defect Migration ........................................................................... 132
   4.1.2. Monte Carlo Methods for Simulation of Interstitial Impurity Migration ................................................. 125
   4.1.3. Macroscopic Analysis of Point Defect Migration ........................................................................... 143
       4.1.3.1. Transport Equation for the Number Density of Point Defects ......................................................... 143
       4.1.3.2. Moment Analysis of the Transport Equation ............................................................................... 145
   4.1.4. Impurity Migration Under Constant Force .................................................................................. 148
       4.1.4.1. Analytical Results ....................................................................................................................... 148
       4.1.4.2. Computational Results ............................................................................................................... 151
   4.1.5. Extension to Realistic Systems and Inhomogeneous Forces .......................................................... 158

4.2. Migration of Interstitial Oxygen in the Dislocation Field of a Silicon Lattice ........................................... 162
   4.2.1. Analysis of Oxygen Diffusion Around a 60° Dislocation in Silicon ......................................... 162
   4.2.2. Constitutive Model for the Drift Velocity and the Diffusivity ................................................... 164
   4.2.3. Diffusivity Data for Oxygen Diffusion in the (110) Plane ....................................................... 167

4.3. Oxygen Concentration Field Around Dislocations in Silicon ....................................................... 168
   4.3.1. Mathematical Formulation ........................................................................................................... 170
       4.3.1.1. Transport Equation and Boundary Conditions ........................................................................... 172
       4.3.1.2. Constitutive Equations for Isotropic and Anisotropic Migration ........................................... 175
   4.3.2. Finite-Element Analysis ................................................................................................................. 177
   4.3.3. Oxygen Atmosphere Formation .................................................................................................. 180

4.4. Summary and Discussion .................................................................................................................. 184
5. EFFECTS OF INTERSTITIAL OXYGEN ON THE MECHANICAL STRENGTH OF SILICON ................................................................. 188

   Dislocation Locking and Unlocking .............................................. 189
   5.1.1. Dislocation Drag in High-Purity Silicon ............................... 190
   5.1.2. Dislocation Drag from Interaction with Oxygen Atmospheres .... 191
   5.1.3. Dislocation Drag from Oxygen Precipitates .......................... 199
   5.1.4. Modeling of Dislocation Locking and Unlocking ................... 205

5.2. Constitutive Model for the Initial Stages of Deformation .............. 210
   5.2.1. Constitutive Model for the Effective Stress ....................... 210
   5.2.2. Constitutive Model of Dislocation Multiplication .................. 217

5.3. Modeling of Plastic Deformation Dynamics .................................. 221
   5.3.1. Simulation of Dynamic Deformation Experiments ................... 221
   5.3.2. Estimation of the Critical Resolved Shear Stress .................. 233
   5.3.3. Limiting Cases of the Constitutive Model of Plastic Deformation Dynamics ................................................................. 237
   5.3.3.1. Large Deformations ..................................................... 237
   5.3.3.2. Small Deformations .................................................... 238

5.4. Summary and Discussion ....................................................... 239

6. DYNAMICS OF INTRINSIC POINT DEFECTS IN SILICON .................... 244

6.1. Vacancies and Self-Interstitials in Silicon ................................ 245
   6.1.1. Structural Configurations and Migration Paths ..................... 245
   6.1.2. Experimental Measurements of Intrinsic Point Defect Properties .... 248
   6.1.3. Self-Diffusion and Dopant Diffusion ................................ 250
   6.1.4. First-Principles Calculations of Intrinsic Point Defect Properties .... 253

6.2. Integrated Atomistic Approach for Calculation of Intrinsic
   Point Defect Properties in Silicon ............................................ 255
   6.2.1. Atomic Model for Crystalline Silicon ................................ 256
6.2.2. Atomistic Simulation Methods ......................................................... 263
  6.2.2.1. Metropolis Importance Sampling and Canonical Ensemble
            Monte Carlo ................................................................. 263
  6.2.2.2. Isothermal-Isobaric Ensemble Metropolis Monte Carlo .............. 266
  6.2.2.3. Simulated Annealling ....................................................... 268
  6.2.2.4. The Steepest Descent Method ............................................. 271

6.2.3. Calculation of the Temperature Dependence of
          Thermodynamic Properties ..................................................... 275

6.2.4. Free Energy Calculations ....................................................... 281

6.2.5. Point Defect Concentrations at Thermal Equilibrium .................... 286

6.2.6. Jump Rate Theory of Point Defect Migration ................................ 289

6.3. Equilibrium and Transport Properties of Intrinsic
          Point Defects in Silicon ....................................................... 291
  6.3.1. Thermodynamic Properties of Perfect Crystalline Silicon ............ 292
  6.3.2. Formation Properties of Extended Point Defects ......................... 299
  6.3.3. Formation Properties of Strictly Localized Point Defects ............. 311
  6.3.4. Thermal Equilibrium Concentrations ....................................... 324
  6.3.5. Diffusion Coefficients and Contributions to Self-Diffusion .......... 328

6.4. Summary and Discussion .................................................................. 341

7. CONCLUSIONS ................................................................................... 344
  7.1. Summary ....................................................................................... 345
  7.2. Directions for Future Research .................................................... 348

REFERENCES ......................................................................................... 353

APPENDIX .............................................................................................. 375
  A.1. Interatomic Forces and Internal Virial .......................................... 375
  A.2. Calculation of the Dynamical Matrix ............................................ 377
LIST OF FIGURES

1.1. Point and line defects in a simple cubic lattice ........................................ 24

1.2. Thermal conductivity of common semiconductor materials vs. best estimates for the critical resolved shear stress (CRSS) at the melting point (Brown, 1988) ........................................ 26

1.3. The Czochralski (CZ) melt crystal growth system (Brown, 1988) .................. 27

1.4. Characteristic length and time scales and types of models used for the analysis of defect dynamics in semiconductor crystals .................................................. 32

1.5. Outline of the integrated methodology for quantitative defect modeling. The broad range of length scales captured by this methodology is shown schematically ........... 33

2.1. A straight edge dislocation in a simple cubic lattice .................................... 38

2.2. A straight screw dislocation in a simple cubic lattice .................................. 39

2.3. The glide plane of a straight dislocation ..................................................... 40

2.4. One configuration of a 60° dislocation in the diamond structure .................. 42

2.5. A typical stress-strain curve for Si exhibiting two yield points. The lower yield stress is usually taken to be equal to the critical resolved shear stress (CRSS) .......... 44

2.6. Configuration of a mobile dislocation in diamond/zincblende structure semiconductor materials ................................................................. 48

2.7. A kinked dislocation line in a crystal with a high Peierls potential energy barrier .......................................................................................... 50

2.8. Dislocation multiplication due to the cross-slip mechanism: (a) formation of a large glide band under the action of stress as a result of dislocation multiplication, and (b) cross-slip of the screw part of a dislocation and initiation of the dislocation multiplication process ........................................................................... 52

2.9. Effects of n-doping and p-doping on dislocation velocities in Ge under high applied stresses (Patel and Chaudhuri, 1966) ...................................................... 55

2.10. Schematic diagram of the dependence of the average dislocation velocity \( v \) on the applied stress \( \tau \) in solute hardened Ge, where the exponent \( m \) in the velocity-stress relation of equation (2.5) is equal to 1 .................................................. 57
2.11. The Portevin-LeChatelier effect in vanadium caused by the interaction between carbon atoms and dislocations (Yoshinaga et al., 1971) ................................................................. 58

2.12. Comparison between (a) an experimentally observed dislocation pattern in a GaAs wafer grown by the LEC method and (b) contours of excess CRSS predicted by simulation (Jordan et al., 1980) .................................................................................................................. 70

3.1. (a) View of a diamond lattice with the (111) planes perpendicular to the plane of the paper. (b) Projection of the diamond lattice onto the (110) plane. The two fcc sublattices and the two types of glide planes (I and II) are shown. (c) Projection of the diamond lattice onto the (111) plane. The Burgers vector $\mathbf{b}$ of the complete dislocation and the Burgers vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ of the two partial dislocations are shown (Alexander, 1986) ......................................................................................................................... 76

3.2. Geometrical interpretation of the notation used in the treatment of random continuous distributions of dislocations ................................................................................................................. 84

3.3. Decomposition of the growing crystal into an interfacial region and a bulk region where dislocation dynamics is analyzed ............................................................................................................. 87

3.4. Dependence of the local dislocation density on local strain after compression creep of germanium samples at $T = 855$ K under stresses of (a) $0.25$ kg/mm$^2$ and (b) $0.5$ kg/mm$^2$ (Alexander and Haasen, 1968) ................................................................................................................................. 103

3.5. Variation of the dislocation density with the distance from the melt/crystal interface predicted by equations (3.62a) and (3.62b). Results are for InP at a constant von Mises stress of $0.1$ MPa. The solid (---) and dashed (- - - -) curves correspond to $m = 1$ and $m = 3$, respectively. The above curves provide upper and lower bounds for the axial variation of the dislocation density for Si ($m = 1.1$), GaAs ($m = 1.7$) and InP ($m = 1.4$) .................................................................................................................................................. 106

3.6. Comparison between the asymptotic results of eqs. (3.62a) and (3.62b) and the numerical solution of eqs. (3.67) and (3.68) for InP in the cases of a constant or decreasing von Mises stress with the distance above the melt/crystal interface: $\sigma_0 = 0.1$ MPa ..................................................................................................................................... 112
3.7. Variation of the plastic strain with the distance from the melt/crystal interface. The solid line is the numerical solution of eqs. (3.67) and (3.68) using a constant temperature gradient of -50 K/cm and a linear axial variation of the von Mises stress from 0.1 MPa at the interface to 1 MPa at a distance from the interface equal to the encapsulant height. The curve (---) is the asymptotic result of eq. (3.70) .................................................. 114

3.8. Axial variation of the dislocation density for a linear axial variation of the von Mises stress with minimum value \( \sigma_0 = 0.1 \) MPa and maximum value \( \sigma_{max} = 1 \) MPa. The solid curve (-----) is the asymptotic result eq. (3.76) and the two curves are the results of numerical solution of eqs. (3.67) and (3.68) for a constant temperature gradient (---) and the fin approximation (---) to the temperature profile .................................................. 115

3.9. Comparison between asymptotic (curves) and numerical results (○, □) for analysis of dislocation multiplication in LEC grown InP with shallow and deep encapsulant layer respectively, based on the model of Alexander and Haasen ........................................ 119

3.10. Schematic diagram of the radial profile of (a) the von Mises stress and (b) the dislocation density in crystals of InP or GaAs grown by the LEC method ................. 121

3.11. Decomposition of the growing crystal into three regions according to the dominant mechanisms of defect dynamics. Dislocation generation dominates in the melt/crystal interfacial region, dislocation multiplication dominates in the dislocation density boundary layer and impurity gettering to dislocation cores dominates in the bulk away from the boundary layer ................................................................. 128

4.1. Outline of the methodology for the theoretical analysis of interstitial impurity migration ............................................................................................................. 133

4.2. Diagram of one-dimensional potential energy \( \varphi \) with respect to the lattice positions \( r \) for the cases of (a) zero, (b) constant and (c) inhomogeneous interaction force \( F(r) \) ........................................................................................................... 141

4.3. Convergence of the components \( D_{ij} \) to a final equilibrium value as a function of time. The time evolution is given by the number of subtrajectories, each equal to 1000 MC steps. The temperature for the simulation is \( T = 1000 \) K. Results are shown for (a) a two-dimensional square lattice and (b) a simple cubic lattice ........................................... 153
4.4. Time dependence of the components of the first moment for diffusion with constant force in a square lattice for \( F_x = -0.5 \, Q^o/a \) and \( F_y = -0.2 \, Q^o/a \): (a) \( T = 1000 \, K \) and (b) \( T = 300 \, K \) ................................................................. 155

4.5. Time dependence of the components of the second central moment for diffusion with constant force in a square lattice for \( F_x = -0.5 \, Q^o/a \) and \( F_y = -0.2 \, Q^o/a \): (a) \( T = 1000 \, K \) and (b) \( T = 300 \, K \) ................................................................. 156

4.6. Convergence of the components \( D_{ij} \) to equilibrium values as a function of time for \( F_x = -0.5 \, Q^o/a \) and \( F_y = -0.2 \, Q^o/a \): (a) \( T = 1000 \, K \) and (b) \( T = 300 \, K \) .............. 157

4.7. Arrhenius plots for the temperature dependence of the diagonal components of the diffusivity tensor: (a) \( F_x = -0.5 \, Q^o/a \) and \( F_y = -0.2 \, Q^o/a \); (b) \( F_x = -0.5 \, Q^o/a \) and \( F_y = 0 \); (c) \( F_x = -0.5 \, Q^o/a \) and \( F_y = 0.5 \, Q^o/a \). The results of the Monte Carlo simulation for \( D_{xx} \) (○) and \( D_{yy} \) (+) are compared with the results of eqs. (49a) (---) and (49b) (- - -). The solid line (———) in each plot gives the two-dimensional random walk result (50) ......................................................................................... 159

4.8. Dependence of the diffusivity components on the magnitude of the component \( F_x \). The results of the Monte Carlo simulation for \( D_{xx} \) (○) and \( D_{yy} \) (△) are compared with the results of eqs. (49a) (---) and (49b) (- - -). Calculations are shown for \( F_y \) equal to 0.1, 0.3, 0.5 and 0.8 \( Q^o/a \) ................................................................. 160

4.9. (a) Schematic representation of a \( 60^\circ \) glide dislocation (line OE) with respect to the conventional cubic unit cell of the diamond lattice. (b) Two-dimensional hexagonal cells of silicon atoms in the \( (110) \) plane (ABCD) that is perpendicular to the dislocation line ............................................................................................................. 163

4.10. (a) The interstitial lattice structure of the two-dimensional silicon lattice in the \( (110) \) plane. (b) The crystallographic axes of the two-dimensional interstitial lattice ............................................................................................................. 165

4.11. Arrhenius plot for the temperature dependence of the diffusivity of interstitial oxygen in silicon in terms of the values of the components of the diffusivity tensor for diffusion in the \( (110) \) plane. Results are determined at the position \( (r = 5a, \, \theta = \pi/2) \) from the center of the dislocation core, as derived from eqs. (60) and (61c) ............................................. 169
4.12. Domains for modelling oxygen field about (a) a regular array of dislocations and (b) an isolated dislocation ................................................................. 174

4.13. Finite element mesh used for the numerical computations of the oxygen field around 60° dislocations ................................................................. 181

4.14. Oxygen distribution around a 60° dislocation in silicon computed using the model of isotropic migration, for $T = 1500$ K, and the dislocation velocities of (a) 0, (b)-(c) $5.5 \times 10^{-4}$ m/s, and (d)-(e) $3.0 \times 10^{-3}$ m/s. The contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 and 10.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$, and (d) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$ ................. 182

4.15. Oxygen distribution around a 60° dislocation computed using the model of anisotropic migration, for $T = 1500$ K, and the dislocation velocities of (a) 0, (b)-(c) $3.0 \times 10^{-4}$ m/s, and (d)-(e) $3.0 \times 10^{-3}$ m/s. The contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 and 10.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$, and (d) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0 and 20.0 of $c_0$ ........................................................................................................ 183

4.16. Oxygen distribution around a 60° dislocation for $T = 920$ K and dislocation velocities of (a) 0.0 m/s, (b) $2.5 \times 10^{-9}$ m/s and (c) $3.0 \times 10^{-8}$ m/s. Contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, 20.0 and 30.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0 and 20.0 of $c_0$, and (c) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 20.0 and 40.0 of $c_0$, respectively ................................................................................................................ 185

4.17. Oxygen distribution around a 60° dislocation for $T = 1073$ K and dislocation velocities of (a) 0, (b) $1.8 \times 10^{-7}$ m/s, and (c) $2.5 \times 10^{-6}$ m/s. The contours shown are (a) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0, 20.0, and 24.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, and 15.0 of $c_0$, and (c) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, 20.0, 25.0, 30.0, and 35.0 of $c_0$ ..................................................................................... 186

5.1. Dislocation velocity as a function of the drag force developed by the interaction between the dislocation and the oxygen distribution; $T = 1500$ K .................. 194
5.2. Comparison of the predictions of the isotropic diffusivity model (○) and the anisotropic diffusivity model (●) for the dependence of the dislocation velocity on the drag force; Results are shown for the temperatures (a) $T = 1500$ K, (b) $T = 1200$ K, (c) $T = 920$ K and (d) $T = 600$ K ................................................................. 196

5.3. Arrhenius plot of the temperature dependence of the maximum value of the dislocation drag force for the isotropic (○) and the anisotropic (●) diffusivity models .......... 197

5.4. Temperature dependence of the critical dislocation velocity predicted from the isotropic (○) and the anisotropic (●) diffusivity models. Cottrell’s approximation (- - - -) also is shown ................................................................. 198

5.5. Dependence of the dislocation velocity (- - - -) on the applied stress for oxygen concentration $c_0 = 1.0 \times 10^{18}$ cm$^{-3}$ at $T = 920$ K. The predicted dependence of the dislocation drag stress on the dislocation velocity (-----) and the stress dependence of the dislocation velocity for pure silicon (- - - -) also are shown ............................................ 200

5.6. Effect of bulk oxygen concentration on the dependence of dislocation velocity on the applied stress for $T = 920$ K ................................................................................................. 201

5.7. Comparison of the predicted dependence of the dislocation velocity on the applied stress neglecting the drag stress due to oxygen inside the core with the experimental data of Imai and Sumino; $T = 920$ K; (a) $c_0 = 7.4 \times 10^{17}$ cm$^{-3}$ and (b) $c_0 = 9.2 \times 10^{17}$ cm$^{-3}$ ........................................................................................................ 202

5.8. Structure of staggered cluster of oxygen in bulk silicon (Needels et al., 1991) .... 204

5.9. Relationship between dislocation velocity and drag stress due to (1) the elastic interaction between the oxygen atmosphere in the lattice and the dislocation ($\tau_{atm}$) and (2) oxygen precipitates inside the core ($\tau_c$). The calculations are for a silicon crystal with oxygen concentration $c_0 = 9.2 \times 10^{17}$ cm$^{-3}$ at $T = 920$ K ............................................. 206

5.10. Comparison between the model predictions for the dependence of the $60^\circ$ dislocation velocity on the applied stress in silicon crystals with oxygen concentrations (a) $c_o = 2.5 \times 10^{17}$ cm$^{-3}$, (b) $c_o = 7.4 \times 10^{17}$ cm$^{-3}$ and (c) $c_o = 9.2 \times 10^{17}$ cm$^{-3}$ at $T = 920$ K (-----) and the experimental data of Imai and Sumino (Δ, ○, □). The dashed line gives the result in high purity silicon ................................................................. 207
5.11. Dislocation velocity as function of the applied stress for various concentrations of interstitial oxygen at (a) $T = 647^\circ C$ and (b) $T = 900^\circ C$. The dashed line gives $v = v(\tau)$ for high purity material .................................................................................................................. 209

5.12. Critical stress $\tau_u$ for dislocation unlocking as a function of oxygen concentration in silicon as a function of temperature ............................................................................................................. 211

5.13. Temperature dependence of the unlocking stress as a function of oxygen concentration ........................................................................................................................................ 212

5.14. Geometrical details for the calculation of the component of the internal stress due to the interaction between a dislocation and a neighboring oxygen cloud. The dislocation glide plane is perpendicular to the $y$ axis .................................................................................................................... 216

5.15. Geometrical details of the initiation of the cross-slip multiplication process due to the formation of a jog of height $z_0$ in a screw dislocation ............................................................................................................. 219

5.16. Comparison between the predicted stress-strain curves in high purity silicon at $T = 800^\circ C$, $\dot{\epsilon} = 1.2 \times 10^{-4} \text{ s}^{-1}$ and $N_0 = 10^4 \text{ cm}^{-2}$, with and without introducing a strain hardening function in the expression for the effective stress, eq. (5.36) .................................................. 224

5.17. Comparison between the (a) predictions of the model, and (b) experimental data for the stress-strain curves at $T = 800^\circ C$, $\dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1}$ and $N_0 = 10^6 \text{ cm}^{-2}$ for several oxygen concentrations .............................................................................................................................. 226

5.18. Dependence of the mobile dislocation density on strain in crystals for several oxygen concentrations during deformation at $T = 800^\circ C$, $\dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1}$ and $N_0 = 10^6 \text{ cm}^{-2}$ ........................................................................................................................................... 227

5.19. Dependence of the upper yield stress on the initial dislocation density $N_0$ with varying oxygen concentrations during deformation at $T = 800^\circ C$ and $\dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1}$. The dashed curves represent two-point fits according to eq. (5.37) .......................................................................................................................................... 228

5.20. Comparison between the predictions of the model and experimental data of the dependence of the upper yield stress on the oxygen concentration during deformation at $T = 800^\circ C$ and $\dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1}$ for two initial dislocation densities ........................................ 230
5.21. Dependence of (a) stress and (b) dislocation density on strain including stages of deformation past the yield region. The concentration of oxygen in the crystal is \( c_0 = 1.5 \times 10^{17} \text{ cm}^{-3} \) and the deformation is carried out at \( T = 900^\circ\text{C} \) and \( \dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1} \). The dashed curve in (b) is a two-point parabolic fit .................................................. 231

5.22. Lower yield stress as a function of oxygen concentration during deformation at \( T = 900^\circ\text{C} \) and \( \dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1} \) ........................................................................................................ 232

5.23. Predictions of the deformation behavior between silicon with oxygen concentration \( c_0 = 1.5 \times 10^{17} \text{ cm}^{-3} \) from the model and of a Maxwell solid with \( t_{rel} = 149 \text{ s} \); \( T = 800^\circ\text{C} \) and \( \dot{\epsilon} = 1.1 \times 10^{-4} \text{ s}^{-1} \) ........................................................................................................ 240

6.1. Self-interstitial configurations in a (110) plane of a silicon lattice. The tetrahedral (T), hexagonal (H) and bond-centered (B) sites and the dumbbell configuration (D) are shown. The arrows denote simple migration paths which involve T and H sites (TH), and B and D sites (BD) and interstitialcy migration paths which involve T and B sites (TB), and T and D sites (TD) ........................................................................................................ 247

6.2. Integrated approach for atomistic calculations of intrinsic point defect properties in silicon ........................................................................................................ 257

6.3. Reduced pair interaction of the interatomic potential of Stillinger and Weber .... 260

6.4. Density dependence of the lattice energy for cubic lattice structures of silicon according to the interatomic potential of Stillinger and Weber ........................................................................ 261

6.5. Convergence of the steepest descent algorithm for the minimization of the Rosenbrock function ........................................................................................................ 273

6.6. Convergence of the steepest descent algorithm for the minimization of the potential energy of a silicon structure slightly perturbed from the ideal diamond lattice configuration. The convergence is demonstrated by (a) the Euclidean norm of the displacements of the atomic coordinates, and (b) the reduced potential energy of the crystal as a function of the number of steepest descent iterations .................................................. 274

6.7. Cubic simulation cell containing 216 silicon atoms in a diamond-cubic lattice configuration ........................................................................................................ 293
6.8. Comparison of the results of cumulant analysis with results from MC simulations for the temperature dependence of the enthalpy of a perfect silicon crystal ................................................. 295
6.9. Comparison of the results for the temperature dependence of the Gibbs free energy of a perfect silicon crystal of cumulant analysis with the results of the quasi-harmonic approximation (QHA), the local-harmonic approximation (LHA), and the temperature-integration scheme of Broughton and Li (1987) ........................................................................ 296
6.10. Comparison of the results of cumulant analysis for $n = 2$ and $n = 3$ with results from MC simulations for the temperature dependence of the density of a perfect silicon crystal ........................................................................................................... 298
6.11. Dependence of the formation energy of vacancies and self-interstitials in Stillinger-Weber silicon on the number of atoms in the simulation cell ................................................. 301
6.12. Temperature dependence of the enthalpy for a silicon crystal with (a) an extended vacancy, and (b) an extended self-interstitial per simulation cell ......................................................... 302
6.13. Local energy minima that correspond to the atomic configuration at the end of a cooling stage as a function of the temperature of the cooling stage using a variation of the simulated annealing method for energy minimization of a silicon crystal that contains (a) a vacancy and (b) a self-interstitial per simulation cell ......................................................... 304
6.14. Atomic relaxation in units of lattice parameter (a) within neighboring shells of (a) an extended vacancy, and (b) an extended self-interstitial in a silicon crystal ...................... 305
6.15. Temperature dependence of the enthalpy ($H_N^L$), Gibbs free energy ($G_N^L$), and vibrational entropy ($S_N^L$) of formation of extended vacancies ($X = V$) and self-interstitials ($X = I$) in silicon ......................................................................................................................... 307
6.16. Comparison of the results of cumulant analysis for $n = 2$ and $n = 3$ with results from MC simulations for the temperature dependence of the density $\rho = \rho(T)$, of a silicon crystal that contains extended vacancies or extended self-interstitials ............................ 309
6.17. Temperature dependence of the relaxation volume ($\Delta V_{rel}^L$) and the formation volume ($V_N^L$) of the extended vacancy ($X = V$) and the extended self-interstitial ($X = I$) ...... 310
6.18. Atomic relaxation in units of lattice parameter \( a \) within neighboring shells of \( a \) a vacancy, and \( b \) a split-vacancy in a silicon crystal ................................................................. 315

6.19. Atomic relaxation in units of lattice parameter \( a \) within neighboring shells of \( a \) a tetrahedral, and \( b \) a hexagonal self-interstitial in a silicon crystal ............................... 316

6.20. Atomic relaxation in units of lattice parameter \( a \) within neighboring shells of \( a \) a bond-centered, and \( b \) a dumbbell self-interstitial in a silicon crystal ......................................................... 317

6.21. Temperature dependence of the enthalpy \( H_X^f \), Gibbs free energy \( G_X^f \), and vibrational entropy of formation \( S_X^f \) of the vacancy \( X = V \) and the split-vacancy \( X = SV \) in silicon ......................................................................................... 319

6.22. Temperature dependence of the enthalpy \( H_X^f \), Gibbs free energy \( G_X^f \), and vibrational entropy \( S_X^f \) of formation of the tetrahedral \( X = T \) and the hexagonal \( X = H \) self-interstitial in silicon ........................................................................................................ 320

6.23. Temperature dependence of the enthalpy \( H_X^f \) and the Gibbs free energy \( G_X^f \) of formation of the bond-centered \( X = B \) and the dumbell \( X = D \) self-interstitial in silicon .................................................................................................................. 321

6.24. Temperature dependence of the formation volume of the vacancy and the extended vacancy ................................................................................................................................. 322

6.25. Temperature dependence of the formation volume of strictly localized self-interstitial configurations ........................................................................................................ 323

6.26. Temperature dependence of the thermal equilibrium concentration of \( a \) vacancies, and \( b \) self-interstitials. The simulation predictions are compared with results from analyses of positron annihilation and dopant diffusion data ........................................ 326

6.27. Atomic jump toward a vacant lattice site. The unrelaxed atomic configuration shows the migration path. The variation of the energy as a function of the migration path coordinate \( \eta \) is shown following two different energy minimization techniques ........... 330

6.28. Temperature dependence of the vacancy diffusion coefficient ........................................ 332
6.29. Temperature dependence of the component of the self-diffusion coefficient due to the vacancy mechanism. Comparison between simulation results and analyses of gold diffusion data ................................................................. 334

6.30. Energy variation as a function of the migration path coordinate $\eta$ for the simple interstitial path TH ................................................................. 335

6.31. Temperature dependence of the self-interstitial diffusivity for the simple interstitial path TH. Comparison of simulation results with experimental data ...................... 338

6.32. Temperature dependence of the component of the self-diffusion coefficient due to the interstitialcy mechanism. Comparison between simulation results and analyses of gold diffusion data ................................................................. 340

6.33. Phase diagram of silicon at zero pressure from MC simulations using the SW interatomic potential including the perfect crystal, the amorphous phase, and crystals that contain vacancies and self-interstitials at different concentrations .......................... 343
LIST OF TABLES

3.1. Parameters of the model of Alexander and Haasen ............................................. 93

6.1. Parameters of the Stillinger-Weber interatomic potential ........................................ 262

6.2. Comparison of the predictions for the Gibbs free energy of the perfect crystal using the QHA, the LHA, and the cumulant analysis with the predictions using the temperature-integration scheme of Broughton and Li (1987) .................................................. 299

6.3. Formation energies and first-neighbor relaxations of intrinsic point defects in silicon predicted by our simulations using the SW interatomic potential and by Baskes et al. (1989) using the modified embedded-atom method (MEAM) ............................................. 313

6.4. Formation properties of strictly localized point defect configurations computed at the reference state, $p = 0$ and $T = 500$ K ............................................................... 318
1. INTRODUCTION

Single semiconductor crystals provide the substrates for electronic and optoelectronic device fabrication technologies. Large crystals of silicon and III-V compound semiconductors, such as gallium arsenide and indium phosphide, are produced by a variety of bulk melt growth techniques (Brown, 1988) where the single crystal is grown from the melt. The single crystals are subsequently cut into wafers which undergo a series of thermal and chemical treatments (Ghandi, 1983; Ravi, 1981) to provide the substrates for micrpelectronic devices, such as light emitting diodes and field-effect transistors. Device performance is governed by the quality of the substrate which depends on the composition and the structural perfection of the semiconductor crystal. The type and density of lattice imperfections and impurities in the semiconductor substrate determine the optical, electronic and mechanical properties of the material (Ghandi, 1983; Ravi, 1981). Very large scale integration and the needed miniaturization of devices requires a deep understanding of the complicated physics of defects in the substrate in order to provide quantitatively accurate predictions of the effects of crystalline defects on device performance.

1.1. Point and Line Defects in Single Crystals

In single crystals the defects of interest are point and line defects. A schematic representation of common point and line defects in crystalline lattices is shown in the two-dimensional projection of a simple cubic lattice in Fig. 1.1. Point defects are either intrinsic point defects or impurity species. The intrinsic point defects are vacancies, which are missing atoms from lattice sites, or self-interstitials, which are extra atoms at interstitial lattice sites (Lannoo and Bourgoin, 1982). Vacancies and self-interstitials are the vehicles of substitutional mass transport in the crystal and govern self-diffusion and substitutional dopant diffusion (Fahey et al., 1989). Most of the $p$–type and $n$–type dopants which are introduced in the crystal to modify the electronic properties of the intrinsic semiconductor are substitutional in the lattice and migrate following mechanisms that involve vacancies and self-interstitials (Fahey et al., 1989). Interstitial impurities, such as oxygen and hydrogen in silicon, also play key roles in the modification of materials properties and migrate by atomic jumps in the interstitial space of the lattice.

Line defects or dislocations in the crystalline lattice are imperfections with the most catastrophic effects on the performance of semiconductor devices. A common line defect, the edge dislocation is shown schematically at the center of the lattice in Fig. 1.1. The effects of dislocations in the substrate on the optical and electronic characteristics of
Figure 1.1. Point and line defects in a simple cubic lattice.
devices, such as Si integrated circuits (IC's), GaAs laser diodes, light emitting diodes (LED's), field-effect transistors (FET's) and InP detectors and light sources have been studied systematically (Jordan et al., 1980, 1981; Jordan and Parsey, 1988). Typical examples of such effects are the reduction of minority carrier lifetimes in Si bipolar transistors with dislocation densities higher than $10^3$ cm$^{-2}$ in the substrate (Jordan et al., 1981) and the decrease in the optical efficiency of LED's with increasing dislocation density in the III-V compound semiconductor substrate, because dislocations act as nonradiative recombination centers (Jordan et al., 1981). The negative effects associated with the presence of dislocations in the semiconductor substrate make the emphasis in the development and improvement of melt growth techniques focus on the production of dislocation free, or at least low dislocation density crystals. However, careful defect engineering during device processing makes use of dislocations as gettering centers for harmful impurities (Kimmerling, 1991).

1.2. Crystal Growth and Defect Generation

The formation of dislocations is associated with the crystal growth process. It has been believed that the main source of dislocation generation is crystallographic glide caused by thermal stresses that exceed a critical stress level, which is referred to as the critical resolved shear stress (CRSS) (Jordan et al., 1980). Such high thermal stresses are produced by the temperature gradients that develop in the crystal during growth (Brown, 1988). Semiconductor materials have been characterized as being either easy or difficult to grow according to the thermal conductivity of the crystal and the magnitude of the CRSS at high temperatures. Materials with low thermal conductivities and low values of CRSS are more difficult to produce, because larger temperature gradients are needed for set heat fluxes and the lower resistance of the lattice to the formation of dislocations leads more easily to dislocation formation. A plot of the thermal conductivity of common semiconductor materials as a function of the CRSS at the melting point is presented in Figure 1.2 (Brown, 1988). According to this criterion, Si crystals are much easier to grow than III-V compound semiconductor crystals; this fact is supported by experimental observations.

The most common methods for the growth of Si and III-V compound semiconductors are the Czochralski (CZ) and the liquid encapsulated Czochralski (LEC) methods (Matlock, 1977; Jordan and Parsey, 1988), respectively. A schematic of the CZ crystal growth system is shown in Figure 1.3. Both CZ and LEC techniques are characterized as meniscus-defined growth methods (Brown, 1988). A seed crystal is dipped into a pool of melt and the thermal environment is varied, so that a crystal grows from the seed as it is
Figure 1.2. Thermal conductivity of common semiconductor materials vs. best estimates for the critical resolved shear stress (CRSS) at the melting point (Brown, 1988).
Figure 1.3. The Czochralski (CZ) melt crystal growth system (Brown, 1988).
slowly raised above the melt pool. The CZ and LEC methods are batchwise processes, where the volume of the melt decreases as the crystal is grown. The Si substrate for most integrated circuits is grown by the CZ technique. LEC growth is used for the production of GaAs and InP. Here, boric oxide (B₂O₃), an inert nonvolatile material with a relatively low melting point, is used to cover the melt to prevent the decomposition of the very volatile melt (Brown, 1988; Jordan and Parsey, 1988). Usually the pressure of the inert gas above the encapsulant layer is maintained significantly above ambient (Völkli and Müller, 1987). There are several advantages of the meniscus-defined methods; these are discussed by Brown (1988). The most important feature of these methods is that the cooling crystal is free to expand and therefore the thermoelastic stresses are not as high as would be produced if the crystal was confined by a surrounding ampoule.

In spite of the similarity of the growth methods, there is a great difference in the quality of Si and II-VI compound semiconductor crystals that are available today. Based on the seed-in and necking procedure invented by Dash (1958a, 1959), the CZ growth of Si has been optimized, so that dislocation free silicon crystals are available in commercial sizes of 10-15 cm diameters and in large quantities (Matlock, 1977; Brown, 1988; Lin and Benson, 1987). Although low dislocation GaAs and InP crystals with diameters smaller than 1.5 cm have been produced by a variety of methods, undoped GaAs and InP boules are not available in commercial sizes with dislocation densities less than 10³ cm⁻² (Jordan and Parsey, 1988). Moreover, the differences in thermal conductivity and the CRSS of the two materials is insufficient to explain the difference in the quality of the materials, which lead to a difference in cost between the materials of a factor of 100 per cm² (Brown, 1988).

The so-called ‘dislocation free silicon’ crystals grown by the CZ method are not perfect. Recent experimental studies (Chikawa, 1987) have proved that Si crystals contain distributions of many types of microdefects. Some microdefects are microscopic dislocation loops formed by the condensation of native lattice defects during growth. The formation of these microdefects starts with the clustering of point defects, either vacancies or self-interstitials. Vacancies and self-interstitials are native lattice defects, which means that they always exist in a crystal at temperatures above absolute zero (Balluffi and Granato, 1979). At each temperature there is an equilibrium concentration of vacancies and self-interstitials which corresponds to a minimum of the free energy of the crystal as a function of point defect concentration (Girifalco, 1973). Point defect concentrations higher than the thermal equilibrium concentrations provide the driving force for the formation of clusters of point defects (Mutaftsiev, 1980). Introduction of vacancies and self-interstitials into the crystalline phase at their equilibrium concentrations at the melting temperature during crystal growth from the melt leads to supersaturation of intrinsic point defects as the crys-
tal cools (Mutaftsiev, 1980) because lowering the temperature leads to significant lowering of the equilibrium concentrations (Balluffi and Granato, 1979). Therefore, temperature gradients associated with the crystal growth process control the formation of clusters of point defects by setting the level of supersaturation of the crystal in point defects. Such clusters transform into faulted dislocation loops after they have reached a critical size (Mutaftsiev, 1980). Microscopic faulted loops in semiconductor crystals with either the diamond (e.g. Si and Ge) or the zincblende (e.g. GaAs and InP) crystal structures are immobile (Alexander and Haasen, 1968). Therefore, they do not undergo multiplication mechanisms that characterize the mobile dislocations under the action of thermal stresses.

Other microdefects are clusters of oxygen atoms (Chikawa, 1987) which is the main impurity in CZ grown Si (Brown, 1988; Sumino, 1987a). The effects of microdefects on the optical and electronic characteristics of Si devices are less harmful than the effects of dislocations (Chikawa, 1987). The most catastrophic effect is that microdefects provide sources for dislocations in later stages of material processing (De Kock et al., 1975).

The most important factor that affects the formation of dislocations in III-V compound semiconductor materials is the nonstoichiometry of the melt (Fujimoto, 1987). This nonstoichiometry introduces a large number of interstitials in the crystal at the melt/crystal interface. Condensation of interstitials in these materials is a major mechanism of dislocation formation. For GaAs growth, the ambient As vapor pressure determines the stoichiometry of the melt and, therefore, the defect densities in the solid (Tomizawa et al., 1987).

Two methods are generally used for the reduction of dislocation densities during growth (Müller et al., 1985). The first is to decrease the temperature gradients in the crystal. This can be accomplished by the development of the appropriate thermal boundary conditions, which depend on several parameters such as the ambient temperature, the pull velocity and the encapsulant height and transparency (Motakef and Witt, 1987; Thomas, 1988; Völkl and Müller, 1989). However, high ambient temperatures for the reduction of temperature gradients in GaAs crystals tend to generate Ga droplets on the crystal surface due to As escape (Tomizawa et al., 1987). The second method is to introduce dopants in the crystal (Sumino et al., 1980; Guruswamy et al., 1987). Some dopants cause solid solution lattice hardening because they retard the motion of dislocations in the crystal (Guruswamy et al., 1987). Another positive effect is that dopants precipitate on dislocations when dislocations move very slowly and cause dislocation locking, so that very high stresses are required to make dislocations move again (Sumino, 1987a). Unfortunately, these dopants affect the electronic properties of the material (Jordan and Parsey, 1988).
The use of ever decreasing device sizes in VLSI applications increases the demands on the crystal quality (Jordan and Parsey; 1988) making necessary the quantitative prediction of crystalline defects. Only substrate materials with very low dislocation densities have the appropriate mechanical and electronic properties for these small device sizes. The goal of defect modeling is the establishment of a framework for quantitatively predicting the effect of operating conditions on the density and type of crystallographic defects, so that strategies to lower this density can be developed.

1.3. Modeling Philosophy

The focus of this thesis is the theoretical analysis and computational study of the dynamics of intrinsic point defects, impurities, and dislocations in single semiconductor crystals produced by bulk melt crystal growth methods. Analysis of defect dynamics during crystal growth or microelectronic device processing involves multiple length and time scales. The migration of intrinsic point defects and impurities is characterized by length scales $O(10^{-10}$ m) and time scales $O(10^{-13}$ s) set by the lattice parameter and the frequency of lattice vibrations, respectively (Flynn, 1972; Girifalco, 1973; Porter and Easterling, 1981). Interactions between mobile dislocations or between point defects and dislocations in semiconductor crystals during deformation at high temperatures are characterized by length scales $O(10^{-6}$ m) and time scales $O(10^{-5}$ s) to $O(10^{-2}$ s). These scales are determined by the distances between dislocations for typical dislocation densities (Sumino, 1987a; Jordan and Parsey, 1988) and by dislocation velocities (Alexander and Haasen, 1968; Sumino, 1987a, 1987b). The objective of the analysis is the prediction of the macroscopic response of a crystal with dimensions $O(10^{-1}$ m) for time scales $O(10^{3}$ s). These time scales are set by growth velocities (Brown, 1988) or strain rates during deformation (Alexander and Haasen, 1968; Sumino, 1987a). The multiplicity of length and time scales associated with defect dynamics in single semiconductor crystals implies that only an integrated methodology for the analysis of thermomechanical and transport processes from the atomic to the continuum scale is appropriate for quantitative defect modeling.

Atomistic simulations using empirical or semi-empirical interatomic potentials (Allen and Tildesley, 1989) are well suited for the study of point defect formation and dynamics (Agullo-Lopez et al., 1988) and the structure, energetics, and stability of point defect clusters of dimensions $O(10^{-9}$ m) to $O(10^{-8}$ m) (Bullough and Perrin, 1968; Sabochick et al., 1988). In addition, atomic modeling is required for the understanding of the structure and the calculation of the energy of the dislocation core as well as the analysis of the atomic rearrangements which result in dislocation motion (Duesbery, 1989). The size of
the dislocation core is $O(a)$, where $a$ is the lattice parameter (Hirth and Lothe, 1982; Duesbery, 1989). The simultaneous microscopic and macroscopic nature of a dislocation is realized by comparison of the size of the dislocation core to the length of a straight dislocation which is set by the characteristic length of the crystal, typically $O(10^{-1} \text{ m})$. Because dislocations do not end inside the bulk crystal (Hirth and Lothe, 1982). Away from the dislocation core the material is accurately described as an elastic continuum and the interactions between dislocations and between point defects and dislocations are analyzed as interactions between stress fields created by the defects in an elastic matrix (Hirth and Lothe, 1982). These interactions are governed by the distances between defects which set the third scale involved in the analysis of defect dynamics. This is a mesoscopic scale intermediate to the atomic and continuum scales. The characteristic length and time scales with the corresponding types of models for for quantitative studies are demonstrated in Fig. 1.4.

The macroscopic response of a solid under the action of a stress field or a nonuniform temperature field is studied using the well-established classical continuum mechanical formalisms of linear elasticity and ideal plasticity theory (Love, 1927; Boley and Wiener, 1960). This macroscopic response is driven by the response of lattice imperfections to external impulses under the simultaneous action of internal fields created by the interactions between defects. Therefore, links are required between the microscopic structure and dynamics of lattice imperfections and impurities and the macroscopic response of the material under the thermomechanical conditions of processing. Such links are established by the development of constitutive theories that relate micromechanical properties, such as dislocation mobility due to microscopic rearrangements of the dislocation line, to the stress field and the temperature history of the material.

The integrated approach for quantitative study of defect dynamics that is followed in this thesis is based on a modeling strategy that combines atomistic simulations, development of constitutive theories, and continuum mechanical analyses. An outline of this approach is given by the flow chart of Fig. 1.5. The starting point is the lattice-scale study of defect dynamics based on atomistic simulations (Allen and Tildesley, 1989) and dislocation theory (Hirth and Lothe, 1982). The results of the microscopic studies are used as input in the development of constitutive equations that relate micromechanical properties to the temperature and stress field in the crystal. In many cases results of first-principle or ab initio calculations (Pantelides, 1990) are incorporated as input into the model. Such constitutive models are coupled with macroscopic transport equations and kinetic models for the propagation and multiplication of defects to calculate defect density profiles in the crystal which are used for the calculation of material properties. The final step is the
Figure 1.4. Characteristic length and time scales and types of models used for the analysis of defect dynamics in semiconductor crystals.
Figure 1.5. Outline of the integrated methodology for quantitative defect modeling. The broad range of length scales captured by this methodology is shown schematically.
comparison with experimental data which also tests the validity of assumptions which are introduced in the models.

1.4. Thesis Objectives and Outline

The objective of this thesis is the systematic theoretical and computational study of the dynamics of point and line defects in semiconductor crystals of industrial interest. The dynamics of mobile dislocations is studied in semiconductor crystals during bulk melt growth to provide thermomechanical criteria for the reduction of dislocation densities in the crystal based on the analysis of the effects of processing conditions on the dislocation density distribution in the growing crystal. The interactions between impurity atoms and dislocations also are analyzed to determine the effects of impurities on dislocation motion and consequently on the mechanical strength of semiconductor crystals. The system of interstitial oxygen in silicon is studied in detail. A systematic theory for the effects of oxygen on the deformation behavior of silicon is developed for the analysis of dislocation retardation and immobilization and the dependence of the mechanical strength of silicon on the oxygen content in the crystal and the deformation conditions. Finally, the dynamics of intrinsic point defects in silicon is analyzed for the prediction of crucial properties for defect engineering in silicon. An integrated computational method is developed based on atomistic simulations for internally-consistent calculations of structural, thermodynamic and transport properties of vacancies and self-interstitials in silicon. This computational approach also provides a robust framework for the study of microdefect and dislocation formation in silicon crystals from the nucleation and growth of clusters of intrinsic point defects.

A review of our knowledge on dislocation formation and dynamics in semiconductor crystals is presented in Chapter 2. A systematic analysis of dislocation multiplication during the bulk melt growth of silicon and III-V semiconductors is carried out in Chapter 3 based on the phenomenological model of Alexander and Haasen for the dynamics of plastic deformation in the diamond/zincblende structure. The migration of interstitial impurities in cubic-type lattices is systematically analyzed in Chapter 4 and a constitutive model is developed for the migration of oxygen in the stress field created by dislocations in silicon crystals. This model is coupled with the macroscopic transport equation for oxygen conservation to calculate the oxygen distribution around dislocations in silicon. The effects of oxygen on dislocation motion is analyzed in Chapter 5 and a model is derived for dislocation drag due to oxygen which predicts dislocation locking and unlocking. Constitutive models are developed for lattice hardening due to oxygen and the effects of
oxygen on dislocation multiplication. These models are used for a complete study of the effects of oxygen on the deformation behavior of silicon. An integrated approach for the study of intrinsic point defect dynamics in silicon based on atomistic simulations is presented in Chapter 6. Formation and migration of vacancies and self-interstitials is studied in detail and the role of these intrinsic defects in self-diffusion and substitutional dopant diffusion is analyzed.
2. DISLOCATION FORMATION AND DYNAMICS IN SEMICONDUCTOR CRYSTALS

The role of dislocations in crystal plasticity and the mechanisms of formation, propagation, and multiplication of dislocations in the diamond/zincblende structure are reviewed in this chapter. Analysis of these mechanisms as well as the interactions between impurities and dislocations is necessary for quantitative modeling of dislocation and impurity dynamics during crystal growth and device processing. The fundamentals of crystallographic glide and the role of dislocations in determining the plastic response of crystals under stress are presented in Section 2.1. The most important mobile dislocation configurations for diamond/zincblende structure crystals are presented and the mechanisms of dislocation motion and multiplication are reviewed in Section 2.2. The interactions between impurities and dislocations and their effects on dislocation mobility are described in Section 2.3. The mechanisms of dislocation generation associated with the crystal growth process are analyzed in Section 2.4; here questions are posed to determine the dominant mechanism for dislocation generation and the identification of dislocation sources independent of the deformation conditions of the crystal. A review also is presented of the experimental evidence for the possible dislocation sources. Finally, attempts of dislocation modeling during crystal growth based on the analysis of the thermoelastic stress field are reviewed in Section 2.5. This modeling is only qualitative, but motivates detailed stress analyses which are necessary for the quantitative treatment of dislocation dynamics in growing crystals.

2.1. Dislocations and Mechanical Strength of Single Crystals

2.1.1. Crystallographic Glide and Dislocations

Plastic deformation is an important characteristic of the mechanical behavior of single crystals. The mechanisms which accompany plastic flow are based on the phenomena of crystallographic glide and mechanical twinning (Schmid and Boas, 1968). Crystallographic glide consists of the slipping of a part of the crystal as a unit across an adjacent part. The surface on which glide takes place is a crystallographic slip or glide plane, the direction of
the motion is the slip or glide direction and the visible intersection of a slip plane with the outer surface of the crystal is a slip band (Kittel, 1968). Glide occurs along crystallographic planes of low Miller indices in the direction of densely packed atomic rows. Consequently, neither the plane nor the direction of slip is determined by the loading but it is fixed by the lattice structure (Schmid and Boas, 1968).

The modern interpretation of the plastic properties of crystalline solids is the theory of dislocations (Hirth and Lothe, 1982; Nabarro, 1987). The idea that crystallographic glide propagates over the slip planes by the motion of dislocations was introduced by Orowan, Taylor and Polanyi (Kittel, 1968). The simplest dislocation configurations are the straight edge and screw dislocations shown schematically in Figures 2.1. and 2.2., respectively, for a simple cubic lattice. The most generic form of a dislocation pattern in a crystal is described by any, not necessarily planar closed curve within the crystal, or an open curve terminating on the crystal surface at both ends (Kittel, 1968).

A dislocation creates a strain field in the crystal, expressed by the strain tensor $\varepsilon$. The displacement vector $u$ of each point of the medium due to the strain field is related to $\varepsilon$ by the equation

$$\varepsilon \equiv \frac{1}{2} [\nabla u + (\nabla u)^T], \quad (2.1)$$

where the notation $(\cdot)^T$ denotes the transpose of a matrix. The strength of a dislocation is determined by the magnitude of the Burgers vector $b$ (Hirth and Lothe, 1982; Nabarro, 1987) defined by the line integral, of the component of $\nabla u$ tangent to a closed curve $c$ that encircles the dislocation:

$$b = - \oint_c \nabla u \cdot dl. \quad (2.2)$$

The Burgers vector is equal to one of the lattice vectors in magnitude and direction. The glide plane of a dislocation is the plane determined by the Burgers vector $b$ and the unit vector $t_d$ that is tangent to the dislocation line; this notation is shown in Fig. 2.3. The Burgers vector of a screw dislocation is parallel to the dislocation line, while the Burgers vector of an edge dislocation is perpendicular to the dislocation line. The line of an edge dislocation is perpendicular to the slip direction. A screw dislocation is parallel to the
Figure 2.1. A straight edge dislocation in a simple cubic lattice.
Figure 2.2. A straight screw dislocation in a simple cubic lattice.
Figure 2.3. The glide plane of a straight dislocation.
slip direction and separates regions of the glide plane where slip has already occurred from regions that have not yet slipped. The most important dislocation configurations in the diamond/zincblende semiconductor crystals are screw dislocations and 60° dislocations. The 60° dislocations have a mixture of edge and screw components, and have the angle between \( \mathbf{b} \) and \( \mathbf{t}_d \) defined as 60° (Alexander and Haasen, 1968); a 60° dislocation in the diamond structure is shown in Figure 2.4. Another common characteristic of imperfect crystals also is portrayed in Figure 2.4. This is the severe distortion of the lattice in the vicinity of the core of the dislocation. Continuum linear elasticity is not valid in this region where the discrete nature of the lattice dominates (Duesbery, 1989). All the experimental methods known for visualizing dislocations, such as revealing etch pits of dislocations penetrating a crystal surface and imaging of dislocations in the crystal interior by infrared light, X-ray topography and transmission electron microscopy, have been applied to diamond structure crystals (Alexander and Haasen, 1968). The etch pit method is accurate for the measurement of dislocation densities up to 10⁸ dislocations/cm² and has been extensively used for the experimental study of dislocations that are formed in a crystal during growth; the so called grown-in dislocations.

2.1.2. Critical Resolved Shear Stress (CRSS)

The force \( \mathbf{F} \) exerted on a dislocation of length \( L \) due to the action of an applied or internal stress field \( \mathbf{g} \) is given by the Peach-Koehler equation (Nabarro, 1987) as

\[
\frac{\mathbf{F}}{L} = (\mathbf{b} \cdot \mathbf{g}) \times \mathbf{t}_d.
\]  

Equation (2.3) indicates that a glide force on the dislocation is produced only by the component of the stress that acts on the glide plane and in the slip direction. Therefore, it is useful to resolve the shear stress into its components in the slip system. Among slip systems with given \( \{h'k'l'\}\{hkl\} \) type, the system with the greatest resolved shear stress dominates in the slip process. The notation \( \{\cdot\} \) denotes equivalent lattice directions and \( \{\cdot\} \) denotes equivalent lattice planes by virtue of the crystal symmetry (Ashcroft and Mermin, 1976); the Miller indices \( h, k \) and \( l \) are defined as a set of integers with no common factors,
Figure 2.4. One configuration of a 60° dislocation in the diamond structure.
inversely proportional to the intercepts of the lattice plane along the crystallographic axes (Ashcroft and Mermin, 1976). In face-centered cubic crystals the dominant slip system is the (1\bar{1}0)\{111\} system, because \{111\} is the family of parallel lattice planes with the shortest interplanar distance and (110) is the direction of closest atomic packing (Kittel, 1968). Resolved shear stresses are frequently used for the calculation of the shear stress on the various slip systems during uniaxial deformation tests (Hirth and Lothe, 1982).

The concept of the critical resolved shear stress (CRSS) in the strength of materials was introduced by Schmid (Schmid and Boas, 1968), who posed the critical shear stress law. Based on experimental studies, a critical value of the resolved shear stress is required for the initiation of crystallographic glide on a substantial scale; the stress necessary for initial glide is the CRSS. Any discussion on the CRSS must be accompanied by the appropriate visualization of the response of the crystal to the stress. This is provided by the stress-strain curve obtained by dynamic deformation experiments carried out under constant temperature and strain rate.

A typical stress-strain curve for Si is presented in Figure 2.5. (Alexander and Haasen, 1968). This figure exhibits two clearly marked phases. In the first (line segment OB) the stress increases linearly with deformation maintained within the elastic limit. The second phase is characterized by substantial plastic deformation at only slightly varying stress. Points B and C are called the upper and lower yield points of the crystal, respectively (Alexander and Haasen, 1968; Yonenaga et al., 1978). The critical resolved shear stress expresses the relationship between the yielding effect and the orientation of the slip systems. In most cases the CRSS is assumed to be equal to the lower yield point stress \(\tau_{ly}\) obtained by uniaxial compression tests and identified as the onset of dislocation formation, which occurs at point A within the linear portion of the stress-strain curve (Müller et al., 1985).

Interesting discussions of the onset of dislocation formation and its relation to the lower yield point are given by Patel and Chaudhuri (1966) in their investigation of charged impurity effects on the deformation of low dislocation density Ge and by Müller et al. (1985) in their study of the deformation behavior of InP crystals. According to the latter study, the onset of dislocation formation occurs at a stress level which lies about 10-20 percent below the lower yield point. An increased formation of dislocations is observed if
Figure 2.5. A typical stress-strain curve for Si exhibiting two yield points. The lower yield stress is usually taken to be equal to the critical resolved shear stress (CRSS).
the stress exceeds the $\tau_{ly}$ value in the initial slope (point A). At the upper yield point B ($\tau = \tau_{uy}$) the sample is filled with dislocations. Therefore, the value of $\tau_{ly}$ appears to be the proper value to characterize the formation of dislocations. However, it is pointed out that $\tau_{ly}$ is the upper limit for the evaluation of CRSS.

2.1.3. Dislocation Motion and Plastic Flow

Plastic flow is studied at the macroscopic level following the principles of continuum plasticity theory for the development of constitutive stress-strain relations necessary (Boley and Weiner, 1960). Significant progress in this field has been made using the analytical framework for the thermodynamics of elastoplastic materials, with internal state variables related to microscopic deformation mechanisms according to continuum dislocation theory (Kratochvil and Dillon, 1969, 1970; Rice, 1975).

The simplest constitutive law that relates the plastic strain field to the dislocation density is the Orowan equation (Gilman, 1968)

$$\dot{\epsilon}^p = b v N_m,$$  \hspace{1cm} (2.4)

where $\epsilon^p$ is a resolved plastic strain on a crystallographic slip system, $N_m$ is the mobile dislocation density, $v$ is the magnitude of the average velocity of the mobile dislocation lines, and the dot denotes differentiation with respect to time. The displacement carried by the dislocations moving on a given glide system is expressed by the magnitude of the Burgers vector $b$. Mura (1963) demonstrated how the scalar constitutive equation (2.4) can be written in general tensorial form, following a rigorous approach established by Kröner (1962; also see Landau and Lifshitz, 1986; Nabarro, 1987).

Equation (2.4) suggests that the plastic response of the crystal is determined by the physics that govern dislocation motion. Dislocation motion is nonconservative, if addition or removal of atoms at the dislocation core is required and conservative otherwise. The glide motion of the dislocation, i.e. the motion on the glide plane, is conservative and is crucial for the crystal plasticity. The climb motion of a dislocation, i.e. the motion perpendicular to the glide plane, is nonconservative and is determined by the interactions between the dislocations and point defects (Balluffi and Granato, 1979).
The Peierls-Nabarro model was the first atomic model for a dislocation in a crystal (Nabarro, 1987). It takes in account the atomic structure in the glide plane, but treats the rest of the lattice as an elastic continuum. In spite of this simplification, the model provides the appropriate physical picture of glide motion as a thermally activated process, where the dislocation overcomes the Peierls barrier. Many mechanisms of dissipation exist during dislocation motion and lead to resistance known as dislocation drag (Nabarro, 1951, 1987; Lothe, 1962; Alshits and Indenbom, 1986); these include the scattering of electrons and phonons and displacement of dissolved atoms by moving dislocations, interactions between the moving dislocation and stationary dislocations, and emission of sound waves caused by the nonuniform dislocation motion.

2.2. Dislocation Dynamics in the Diamond/Zincblende Structure

Analysis of dislocation dynamics in the diamond/zincblende structure focuses on the microscopic mechanisms of dislocation propagation and multiplication and how these mechanisms are affected by the thermomechanical conditions of processing. The structure of the mobile dislocation and the role of irregularities along the dislocation line are of central importance in the development of phenomenological theories and constitutive models that give satisfactory interpretations of experimental observations of dislocation mobility and plastic deformation. Such phenomenological models provide the basis for quantitative modeling of dislocation dynamics and are presented in detail in Chapter 3. This section highlights the underlying mechanisms that govern plastic deformation in the diamond/zincblende structure.

2.2.1. Mobile Dislocation Configuration

An early review on mobile dislocation configurations was given by Alexander and Haasen (1968). According to theoretical estimates of the free energy of various dislocation configurations and experimental observation of dislocations, they concluded that the mobile dislocations in the diamond/zincblende structure are complete dislocations that are not connected to stacking faults. In addition, the experimental work of Schäfer (1967) proved
that the typical dislocation configuration was a loop with the shape of a half hexagon. There are three possible Burgers vectors \( \mathbf{b} \) for this loop; Alexander and Haasen (1968) attempted unsuccessfully to resolve the direction of \( \mathbf{b} \).

This problem was recently solved by the X-ray topographic study of Sumino et al. (Sumino and Harada, 1981; Imai and Sumino, 1983). At high stress levels all moving dislocations that were observed were in the shape of a half hexagon and each segment was straight along the \( (110) \) direction, irrespective of the purity of the crystals. Shape perturbations at lower stresses occur due to the interaction of the dislocation with solute atoms as described in Section 2.3. The Burgers vector was determined to be \( \mathbf{b} = 4a[0\overline{1}1] \), confirming the assumption first made by Chaudhuri et al. (1962). The schematic of the configuration of mobile dislocations in Si and III-V compound semiconductors is shown in Figure 2.6.

2.2.2. Dislocation Mobility in High-Purity Crystals

Measurements of dislocation velocities in high-purity semiconductor crystals have been carried out using the double etch method (Johnston and Gilman, 1959; Patel and Chaudhuri, 1966) and by X-ray topography (Imai and Sumino, 1983). These measurements established a constitutive expression for the average velocity of an isolated dislocation given by

\[
v = B_0 \exp\left(-\frac{Q}{kT}\right)\tau^m,
\]

where \( B_0 \) is a phenomenological mobility coefficient, \( k \) is Boltzmann's constant, \( Q \) is the activation energy for dislocation glide, and \( \tau \) is the applied stress resolved on the glide plane and along the glide direction. Equation (2.5) implies that dislocation motion is a thermally activated process characterized by an energy barrier which is directly related to the Peierls barrier (Nabarro, 1987).

Equation (2.5) also implies that there is a diffusive mechanism associated with dislocation motion in the diamond structure, as noted by the thermally activated process expressed by the Boltzmann factor. The generally accepted theory for the interpretation of the Boltzmann factor is the theory of double kink formation, originally introduced by
Figure 2.6. Configuration of a mobile dislocation in diamond/zincblende structure semiconductor materials.
Lothe and Hirth (1959). Kinks are steps in the dislocation line in its glide plane, which exist because of the Peierls potential energy, so that the dislocation line remains in the minima of the potential, called the Peierls valleys. A straight length of dislocation lying in a Peierls valley has a thermal equilibrium population of kink pairs expressed as

$$c_k \propto \exp(-\frac{G_k}{kT}),$$

where $G_k$ is the free energy for the formation of a kink. A kinked dislocation line is shown in Figure 2.7. Calculations of $G_k$ for diamond structure crystals characterized by a high Peierls barrier (Alexander and Haasen, 1968) give values in good agreement with the experimental values of $Q$ in equation (2.5). The agreement becomes excellent if one adds the activation energy for kink diffusion to $G_k$ (Imai and Sumino, 1983). Kink diffusion occurs by thermal fluctuations in the radiation force on the kink, caused by the scattering of sound waves by the kink. An approximate calculation of the radiation force was carried out by Lothe and Hirth (1959). When the kink moves, it runs into more sound waves on one side than on the other establishing a net radiation force, which determines the mobility of the kink according to Einstein’s equation. Lothe and Hirth (1959) developed a model for the kink flux that included terms for diffusion and drift of kinks and carried out the first calculation of the velocity of an advancing dislocation in its glide plane. This expression implies that dislocation motion in materials with high Peierls barriers is characterized by diffusive rearrangements governed by the dynamics of kink pairs.

Further study of kink dynamics requires lattice-scale modeling to derive a micromechanical model of dislocation motion. The first attempt at this style calculation was presented by Kawata and Ishioka (1985), who carried out a Monte Carlo simulation of the stochastic process of kink motion without including the interaction energies between kinks. A lattice-scale study of kink dynamics based on atomistic simulations has not been attempted in this thesis for reasons that will become apparent in Chapter 3. Dynamics of isolated dislocations in high purity semiconductors is described by equation (2.5), which is extended in the analysis of Chapter 5 to capture the effects of impurities on dislocation motion.
Figure 2.7. A kinked dislocation line in a crystal with a high Peierls potential energy barrier.
2.2.3. Dislocation Multiplication

The mechanism of multiplication in the diamond structure must be analyzed for the development of a constitutive model for the rate of dislocation production due to self multiplication. Recent careful observations of dislocation multiplication in Si have been reported by Sumino and Harada (1981). A first useful observation is that the widths of slip bands increase with time, suggesting that dislocations transfer from one slip plane to another. A direct observation of the major multiplication process in a bulk crystal is available (Sumino and Harada, 1981). It takes place when a screw component of a dislocation with \( \mathbf{b} = \frac{\alpha}{2}[0\bar{1}1] \), which moves on a \((\bar{1}11)\) plane from the \((111)\) surface to the \((\bar{1}1\bar{1})\) surface, crosses a 60° component of a dislocation with \( \mathbf{b} = \frac{\alpha}{2}[10\bar{1}] \) that moves on a \((1\bar{1}1)\) plane from the \((\bar{1}1\bar{1})\) surface to the \((111)\) surface. It has been clearly seen that the multiplication proceeds in the \((\bar{1}11)\) plane (Sumino and Harada, 1981). The Burgers vector of the newly created dislocation is \( \frac{\alpha}{2}[0\bar{1}1] \). The multiplication has been interpreted as the cross-slip of the screw part of the first dislocation onto the \((111)\) plane. The screw component acts thereafter as a pole (Nabarro, 1987) for the multiplication of the first dislocation.

The cross-slip mechanism for multiplication was experimentally observed in LiF by Johnston and Gilman (1959, 1960) and is now believed to be the principal multiplication mechanism for this material (Nabarro, 1987). Johnston and Gilman (1959) introduced a kinetic expression for dislocation multiplication due to the cross-slip mechanism as

\[
\dot{N} = \beta v N,
\]  

(2.7)

where \( N \) is the density of mobile dislocations in the crystal.

The mechanism of dislocation multiplication is based on the expansion of an isolated dislocation loop under the action of stress and the formation of new loops within the original loop, which tend to multiply in the same manner to form a large glide band. This mechanism is presented in Figure 2.8a. The details of the mechanism are given in Figure 2.8b. Segments of the dislocation line, which are of pure screw type, move conservatively in any plane which contains \( \mathbf{b} \). In a crystal the screw segments can often move easily only in planes which are crystallographically equivalent to the glide plane. The deviation of a
Figure 2.8. Dislocation multiplication due to the cross-slip mechanism: (a) formation of a large glide band under the action of stress as a result of dislocation multiplication, and (b) cross-slip of the screw part of a dislocation and initiation of the dislocation multiplication process.
screw segment from a glide plane into one of these equivalent planes is known as cross-slip. If the points B and B' of the screw component in the cross-slip plane become fixed, the segment BB' acts as a Frank-Read source (Frank and Read, 1950) in the new glide plane. The screw component in the cross-slip plane becomes curved when a stress is applied. The curved dislocation segment expands under the action of stress as demonstrated in Figure 2.8b and the points B and B' are held fixed. Two portions of the curved dislocation segment are annihilated when they meet. The next configuration is a closed loop which continues to expand. A linear dislocation segment that joins B and B' is formed and starts the process again. This multiplication process was first analyzed by Frank and Read (1950).

The observations of Sumino and Harada (1981) provide considerable evidence that the cross-slip mechanism is the principal multiplication mechanism in the diamond structure. In addition, their X-ray topographic study showed Frank-Read sources in the interior of the crystal. A similar observation was made for the first time by Dash (1956) in Si by chemical etching and dislocation decoration. Finally, it should be mentioned that both the observations of Sumino and Harada (1981) in silicon and the observations of Johnston and Gilman (1960) in LiF suggest that the formation of a jog due to dislocation intersection marks the onset of dislocation multiplication. This observation also is supported by the experimental observations of Low and Turkalo (1962). A detailed analysis of dislocation multiplication is presented in Chapter 5 based on jog formation and includes a derivation of a more complete form of eq. (2.7) and a model for including the effects of impurities.

2.3. Effects of Impurities on Dislocation Mobility

The mechanical strength of a crystal is determined by the density of mobile dislocations. This is implied by the Orowan equation (2.4) according to which the plastic strain rate is proportional to the mobile dislocation density in the crystal. Immobilization or retardation of dislocation motion improves significantly the mechanical strength of the crystal. There is strong experimental evidence (Alexander and Haasen, 1968) that introduction of dopants in the crystal affects dislocation mobility. Dopants or impurities
are distinguished from atoms added in stoichiometric quantities for the production of compound or alloy semiconductors. Impurities exist in the melt from dissolution of the crucible material, erosion of the system components, or from the inert gas ambient in CZ or LEC systems and their concentrations are satisfactorily controlled (Brown, 1988).

Alexander and Haasen (1968) reviewed the early experimental work on the effects of doping in diamond structure crystals. Chaudhuri et al. (1962, 1963) used the double etch method (Johnston and Gilman, 1959) to measure dislocation velocities in the diamond structure. Patel and Chaudhuri (1966) were the first to measure the influence of doping on dislocation velocities. They found that dislocation velocities in Ge increased with arsenic doping and decreased with gallium doping if the dopant concentration was above a certain level. The effects of arsenic doping and gallium doping on dislocation velocity in Ge under high applied stress is summarized in Figure 2.9. The conclusion is that dislocation motion in Ge is accelerated by donor doping and retarded by acceptor doping, indicating an electronic nature of the dopant-dislocation interaction.

Sumino (1987a) has given the most recent review of the effects of impurities on dislocation dynamics in silicon crystals. Electrically inactive and acceptor impurities do not affect the dislocation velocity when stresses higher than about 5 MPa are present, while donor impurities enhance the velocity. With stresses lower than 1-2 MPa, impurities in Si generally interrupt the motion of dislocations (Sumino and Imai, 1983; Imai and Sumino, 1983). The mechanical strength of silicon is improved markedly by doping with nitrogen and oxygen at concentrations in the range 1-10 ppm (Yonenaga et al., 1984).

Seki et al. (1976, 1978) reported the growth of dislocation free InP and GaAs by the LEC technique using various dopants with concentrations higher than $10^{18}$ cm$^{-3}$ and ascribed the effect of the dopants to the bond formed between impurity atoms and host crystal atoms. Jacob et al. (1982) gave a review of dislocation density data in GaAs and reported the formation of microdefects, which may lead to dislocation formation in the later stages of processing. Guruswamy et al. (1987) reported high temperature experimental results of the mechanical strength of undoped and In-doped GaAs, grown by the LEC method. These measurements show that indium concentrations of about $10^{20}$ cm$^{-3}$ increase the yield stress by almost a factor of two with respect to the value of the undoped
Figure 2.9. Effects of n-doping and p-doping on dislocation velocities in Ge under high applied stresses (Patel and Chaudhuri, 1966).
crystal. Isoelectronic doping of GaAs with In, not only results to significant solid solution hardening, but also causes a minimal influence in its electronic behavior. A systematic experimental study of the mechanical strength of InP was reported by Völkl et al. (1987). The lower yield point \( \tau_{ly} \) of the material was studied as a function of temperature, strain rate and the concentration of a dopant species \( c \), providing the basis for a constitutive relation \( \tau_{ly}(c) \) written in terms of a power law.

The constitutive relation \( \tau_{ly} = \tau_{ly}(c) \) was first reported by Brion et al. (1971), as an interpretation of their experimental results in the deformation behavior of highly doped Ge using the Haasen model. The most interesting point of this study was the verification of a discontinuity in the velocity-stress relation of equation (2.5) at a stress equal to the lower yield stress. This was also observed experimentally (Siethoff, 1970). The curve \( v(\tau) \) is presented in Figure 2.10 for Ge with an exponent \( m \) of the \( v(\tau) \) power law equal to 1. Three regions A, B and C exist, which are characteristic of the dislocation motion in solute hardened materials. At high dislocation velocities in region C, the solute atoms are essentially immobile and the dislocations move over fixed obstacles. At low velocities in region A, solute atoms migrate by diffusion towards the dislocation line, where they are attracted as they generally find energetically favorable sites in the stress field of the dislocation. Therefore, the solute concentration increases near the dislocation, forming a so-called Cottrell cloud (Cottrell, 1953; Nabarro, 1987) and is dragged along with the dislocation. In region B, a transition occurs due to the increased applied stress, so that the dislocation can tear away from the Cottrell cloud and passes to region: C. The existence of this transition region has been used to explain the Portevin-LeChatelier effect of repeated yielding which results in the appearance of serrations after the yield region of the stress-strain curve (Yoshinaga et al., 1971; Nabarro, 1987). The appearance of serrations in the stress-strain of vanadium caused by the interaction between carbon atoms and dislocations (Yoshinaga et al., 1971) is demonstrated in Fig. 2.11. However, the analysis in the seminal work of Brion et al. (1971) does not capture the formation of solute aggregates inside the dislocation core. The present understanding of solute solution hardening has been reviewed by Haasen (1979) and Nabarro (1984).

Accumulation of solute atoms in the core regions of dislocations, while dislocations
Figure 2.10. Schematic diagram of the dependence of the average dislocation velocity $v$ on the applied stress $\tau$ in solute hardened Ge, where the exponent $m$ in the velocity-stress relation of equation (2.5) is equal to 1.
Figure 2.11. The Portevin-LeChatelier effect in vanadium caused by the interaction between carbon atoms and dislocations (Yoshinaga et al., 1971).
are kept at rest at elevated temperatures, leads to immobilization of the dislocations under low applied stresses. This effect is called *dislocation locking*. The external stress needed to start a locked dislocation in motion is called the *unlocking stress* and its magnitude has been measured as a function of dopant concentration, temperature and the length of time that the material is kept at high temperature (Sumino and Imai, 1983; Sumino, 1987a). However, systematic theoretical analysis of dislocation locking and unlocking has not been carried out to quantitatively predict the effects of impurity precipitates inside dislocation cores. Only a simple phenomenological treatment of dislocation unlocking as a thermally activated process has appeared in the literature (Sumino and Imai, 1983).

Finally, Jordan and Parsey (1988) beautifully demonstrated the formation of point defect atmospheres around dislocations and the accumulation of point defects at dislocation cores in LEC grown GaAs crystals. They compared the experimentally measured density distribution of the EL2 defect, an As-antisite (As$_{Ga}$) point defect, with the dislocation density distribution. There is a strong correlation between the two distributions due to the process of formation of the Cottrell cloud (Jordan and Parsey, 1988).

For quantitative analysis of the effects of impurities on dislocation dynamics the potential energy of the interaction between an impurity atom and a dislocation has to be modeled; at least the dominant component of this interaction must be described accurately. The interaction potential between a spherically symmetric point defect and an edge dislocation is taken to be the leading order linear elastic interaction potential due to the size misfit of the point defect in the elastic material matrix (Eshelby, 1976; Hirth and Lothe, 1982; Nabarro, 1987): this potential is

$$U(r) = C_{e} \frac{\sin \theta}{r}, \quad (2.8)$$

where the dislocation lies along the z-axis of a polar cylindrical coordinate system $(r, \theta, z)$. $C_{e}$ is a function of the Burgers vector $b$ and the elastic constants of the material. Equation (2.8) also gives the dominant term of the interaction between a spherically symmetric point defect and a mixed dislocation, i.e. a dislocation with both edge and screw components (Hirth and Lothe, 1982). For a screw dislocation, both linear elastic interactions and the leading order term of the nonlinear elastic interactions (Hirth and Lothe, 1982; Nabarro,
are taken into account and are both expressed by the potential

$$U(r) = \frac{B}{r^2},$$

(2.9)

where $B$ is again a function of $b$, the elastic constants and the Gruneisen constant $\gamma$ of the material (Kittel, 1968). Analogous expressions exist for the elastic interactions between dislocations and nonspherically symmetric point defects (Cochardt et al., 1955).

Equations for the electrostatic interactions (Haasen, 1979; Balluffi and Granato, 1979; Nabarro, 1987), require a detailed quantum mechanical study and are not available, in general, in forms as simple as equations (2.8) and (2.9). Simple closed-form expressions for these interactions have been reviewed by Bullough and Newman (1970) and by Nabarro (1987). The interaction between electrically active impurities and dislocations is long-range and at least of the same order-of-magnitude of the leading order linear elastic interaction due to size misfit (Bullough and Newman, 1970). The electrical activity of dislocations also arises from irregularities on the dislocation line, such as kinks and jogs (Sumino, 1987a). The atomic configuration at the dislocation core is severely different from the configuration in the bulk crystal. Therefore, the electronic structure of an impurity atom in the dislocation core region may be different than the electronic structure in the bulk crystal. Ab initio calculations have begun to answer some of the questions regarding the electrical interactions between dislocations and electrically active dopants in semiconductors (Heggie et al., 1991).

The interactions between impurities and dislocations are commonly used to suppress the warpage of silicon wafers caused by thermal stress during thermal cycling (Sumino, 1987a) and also to getter harmful impurities from electrically active regions of device elements (Kimmerling, 1991). Therefore, understanding of the nature of the impurity-dislocation interaction and predicting its effects on the material properties is valuable in order to find combinations of materials and processing conditions that result in the production of devices with higher and more reliable performance.
2.4. Dislocation Generation During Crystal Growth

The origin of dislocations, in substrates is associated with the crystal growth process. These dislocations are macroscopic line defects and must be distinguished from microscopic dislocation loops, which are present in dislocation free crystals. The number density and size of defects in the crystalline solid depend not only on the growth conditions, such as the growth velocity, the temperature gradient and the ambient temperature, but also on more intrinsic factors, like the stoichiometry of the melt, the impurity content and the structure of the melt/solid interface. These factors are influenced by the growth conditions, which must be controlled to yield the optimum crystal quality.

The seeding and necking procedure invented by Dash (1958a, 1958b, 1959) led to dislocation free Si crystals grown by the CZ technique. The work of Dash has pointed out the importance of both extrinsic and intrinsic factors on crystal perfection. The most important extrinsic factors are the purity of the seed and the thermal environment, while the concentration of vacancies, dependent on the temperature gradient and the growth velocity, and the vacancy-dislocation interactions seem to be the most important intrinsic factors. In spite of the impact of the necking technique in the production of high quality Si crystals, the problem of semiconductor crystal perfection has not been solved for either Si or III-V compounds.

The observation of microscopic dislocation loops (Abe et al., 1983), with harmful but not catastrophic effects in dislocation free Si crystals points out the role of native lattice point defects and atomic size impurities on larger scale defect formation. These defects were invisible using the experimental methods available to Dash. Controlling the thermal conditions near the melt/solid interface of GaAs to be the same as those in growing a dislocation free Si crystal is not enough for growing even a low dislocation density GaAs crystal (Sumino, 1987b). The increased difficulty in the perfection of III-V compound semiconductors lies on the fact that such materials are not monatomic and the two different atomic species must be added in stoichiometric quantities. The nonstoichiometry of the compound is considered as the most important candidate for the high dislocation densities (Fujimoto, 1987) observed in compound semiconductors.
All the possible mechanisms of dislocation formation during growth are presented in Section 2.4.1. The dominant mechanism of dislocation formation is the condensation of point defects. This mechanism is analyzed in some detail in Section 2.4.2 and recent experimental evidence is presented in Section 2.4.3 to support this mechanism.

2.4.1. Mechanisms of Dislocation Formation During Growth

The most recent review of the origin of dislocations in crystal growth was given by Mutaftsiiev (1980), who systematically classified the various mechanisms of dislocation formation during growth from vapor, solution or melt. Crystals grown from the melt are the most imperfect ones. However, dislocations do not intervene in the growth mechanism in melt crystal growth. The spiral or screw dislocation growth mechanism is important only in vapor or solution crystal growth, as implied by the structure of the growing interface (Woodruff, 1973). The three main mechanisms of dislocation formation during melt crystal growth are: (i) the entrapment in the crystal lattice of macroscopic impurity particles or the incoherent impingement of growth fronts, growth particles, or dendrites, (ii) the development of thermal stresses due to the nonuniform temperature distribution in the crystal, and (iii) the formation of vacancy clusters at the growing interface or in the bulk and the transformation of the clusters into dislocation loops.

According to Mutaftsiiev (1980), the last formation mechanism is the only one which must be related to the general mechanisms of crystal growth, because it does not invoke any foreign factors. Factors such as high impurity concentrations in the melt or very high temperature gradients in the solid, are controlled in most crystal growth systems. The control of the concentration of impurities in the melt for high purity products does not permit the formation of aggregates of macroscopic sizes \(O(10^{-6} \text{ m})\) or larger. The incoherent impingement of two lattices is also excluded for the CZ or LEC growth techniques that are characterized by very low growth rates, which do not permit dendritic solidification (Elbaum, 1960). Dendritic patterns are found in multicomponent crystals grown by directional solidification with higher growth rates (Woodruff, 1973). Therefore, the mechanism of dislocation formation due to entrapement of macroscopic impurities or incoherent
impingement of growth fronts is excluded for CZ and LEC growth systems. However, impingement of two lattices is the principal mechanism of dislocation formation in strained layer heterostructures and has been extensively studied by atomistic simulations (Baskes et al., 1988; Dodson and Taylor, 1986; Dodson, 1987).

The mechanism of dislocation formation due to thermal stresses is based on the idea of crystallographic glide and was described in detail in Sections 2.1.1 and 2.1.2. However, the experimental work of Dash (1958b, 1959) has proved that this mechanism is responsible for dislocation multiplication, \textit{but not for dislocation formation}. Dash's experiments showed that plastic deformation in Si crystals due to thermal stresses can occur only if dislocations are already present in the crystal. If a crystal is free from dislocations, even a thermal shock produced by withdrawal and reinsertion of the crystal into the melt bath does not provide a single dislocation.

The experimental work of Dash (1958b, 1959) was extremely important. It led to the development of the necking procedure for the growth of dislocation free silicon crystals by the CZ method and provided a lot of physical insight for the action of imperfections in the growing crystal. According to Dash, the main source of dislocations is imperfections in the seed crystal. These imperfections propagate in the growing crystal and multiply due to thermal stresses. If the initial growth rate is $O(10 \text{ cm/hr})$ the crystal diameter increases rapidly and the produced material has a high dislocation density. Careful treatment of the seed crystal and an increased growth velocity at the early stages of growth provide the method to remove a small of dislocations that possibly exist in the solid at these stages. This is the necking procedure, according to which a growth velocity $O(100 \text{ cm/hr})$ is maintained initially and a long thin neck grows. Any dislocations that may exist in the neck are easy to remove. After the growth of the neck, the growth velocity decreases to $O(10 \text{ cm/hr})$ and dislocation free crystals with large diameters are produced.

The interpretation of the mechanism for removing dislocations was given by Dash (1959) and is based on his experimental observations and the theoretical work of Ham (1959) on point defect diffusion. The equilibrium concentration of vacancies $c_V^q$ in a solid at temperature $T$ is proportional to a Boltzmann factor that contains the formation
enthalpy of vacancies $H_V^f$ (Ballufi and Granato, 1979)

$$c_V^e \propto \exp\left(\frac{-H_V^f}{kT}\right).$$  \hspace{1cm} (2.10)

Analogous relations hold for the equilibrium concentrations of all types of point defects in the crystal (Ballufi and Granato, 1979). A high temperature gradient and a high growth velocity lead to supersaturation of vacancies in the crystal. The temperature effect is obvious from equation (2.10). The effect of the growth velocity is explained in Section 2.4.2. Supersaturation of vacancies leads to climb of edge dislocations, while screw dislocations climb into a helix (Ballufi and Granato, 1979). This climb motion leads to removal of dislocations from the crystal in the thin neck region. In addition, the small diameter of the neck enhances the diffusion of vacancies to the neck surface, as explained in Section 2.4.2 and inhibits the formation of large vacancy clusters.

The above analysis and the experimental evidence presented in Sections 2.4.2 and 2.4.3 restricts the mechanism of dislocation formation in CZ and LEC growth to the condensation of native lattice defects into clusters in the crystal near the melt/solid interface.

2.4.2. Formation of Dislocations by Point Defect Clustering

The idea of dislocation formation caused by vacancy condensation was introduced by Nabarro and Frank (Mutaftsiev, 1980). Vacancies in equilibrium with the lattice at the melting point, supersaturate as the temperature decreases in the crystal and the nucleation of larger holes or voids is favored. The shape of these voids is difficult to determine. The calculation of any equilibrium form invokes thermodynamic properties, such as the surface free energy of the crystal, the values of which are not representative for a cavity of several atoms only. In addition, the cavity shape obtained by growth must depend essentially on kinetic factors, such as the diffusion rate of simple vacancies along certain crystallographic planes. Frank assumed that the cavities are disc-shaped. After reaching some critical diameter, such a disc can collapse and form a stacking fault bordered by a sessile edge dislocation loop (Mutaftsiev, 1980).

Schoeck and Tiller (1960) calculated the minimum radius at which a disc may be
considered to have collapsed and Elbaum (1960) estimated the maximum size of a loop depending on temperature, elastic constants and formation energy of vacancies. The experiments of Howe and Elbaum (1961) on Al growth by the CZ technique confirmed the existence of a critical diameter of the crystal below which vacancy supersaturation is too small to nucleate dislocations. The interpretation for the existence of a critical diameter was based on the solution of the diffusion equation for the vacancies in the crystal. A large crystal diameter maintains a vacancy supersaturation by not permitting vacancies to escape from the interior to the surface of the crystal. High cooling rates typical of Al growth by the CZ method lead to high radial temperature gradients that lower the temperature near the crystal surface significantly with respect to temperatures in the crystal interior. Lower temperatures lead to lower values of the vacancy diffusion coefficient slowing the diffusion of vacancies to the crystal surface. The climb of the resulting sessile edge dislocations and the possibility of their transformation into screw dislocations were discussed by Kuhlmann-Wilsdorf (1958). The common characteristic of the above studies was that the excess vacancy concentration existed only due to crystal cooling during solidification.

Webb (1962) was the first to present a model for the kinetics of trapping a concentration of vacancies, higher than the equilibrium vacancy concentration at the melting point, during growth. Assuming a diffuse melt/solid interface with a thickness \( d \) greater or equal to one atomic layer, Webb concluded that the trapped vacancy concentration depended on the dimensionless parameter \( D_v/Vd \), where \( D_v \) is the vacancy diffusion coefficient in the bulk crystal and \( V \) is the growth velocity. Webb’s analysis was motivated by the mechanism of dislocation formation by condensation of vacancies at the melt/solid interface originally suggested by Teghtsoonian and Chalmers (Mutaftsiev, 1980).

Later, Bolling and Fainstein (1972), reconsidered the nucleation of dislocation loops by condensation of vacancies at the growth interface and presented an elegant thermodynamical model for dislocation loop formation. The importance of their work lies on the experimental evidence for the formation mechanism in Pb. The concentration of vacancies, necessary for the development of Bolling and Fainstein’s thermodynamical model was calculated by an approximate solution of the vacancy transport equation in the interfacial
layer. The result was (Bolling and Fainstein, 1972)

\[ c_q = c_{v}^{eq} + \frac{c_L - c_{v}^{eq}}{1 + \frac{D}{Vd}}, \]  

(2.11)

where \( c_q \) is the vacancy concentration in the interfacial layer, \( c_{v}^{eq} \) is the equilibrium vacancy concentration in the solid at the melting point, \( c_L \) is the concentration of free volume in the liquid, \( D \) is the diffusion coefficient of the vacancies in the solid, \( V \) is the growth velocity and \( d \) is the thickness of the interface layer. The interface thickness \( d \) was calculated using a modification of Temkin's mean field model (Woodruff, 1973) for the structure of the melt/solid interface.

Finally, experimental data for the formation energies and diffusion coefficients of vacancies and self-interstitials in Si (Chikawa, 1987) indicate that condensation of both types of native lattice point defects must be taken into account for the analysis of the dislocation formation mechanism. In addition, as is clarified in the next Section, excess self-interstitials concentrations in GaAs is the basic source for grown-in dislocation formation in LEC growth.

2.4.3. Experimental Evidence for Dislocation Loop Formation in Silicon and III-V Compounds

Microdefects have been observed as swirl patterns in Si crystals grown by the CZ technique or the floating zone (FZ) technique (Abe et al., 1983; Chikawa, 1987). These swirl defects appear in a striated distribution and are classified in two types: A- and B-defects. A-defects were identified by transmission electron microscopy to be interstitial-type dislocation loops and are thought to be formed by supersaturated self-interstitials. B-defects were identified as interstitial-type dislocation loops of smaller size, invisible by transmission electron microscopy unless decorated by Cu (Chikawa, 1987). According to De Kock et al. (1975), dislocations are introduced in the crystal, when the A-defects reach a critical size, above which these defects may undergo self-multiplication due to thermal stresses, following the cross-slip mechanism, as described in Section 2.2.3.

Swirl defects are not the only microdefects observed in Si crystals. The D-defects were
originally observed by Roksnoer et al. (Chikawa, 1987) and are vacancy agglomerates in crystals grown at higher growth velocities. The appearance of both swirl and D-defects led to the hypothesis of the equilibrium coexistence of vacancies and interstitials at high temperatures. The I-defect was originally reported by Abe et al. (1983), who identified these as microprecipitates of oxygen uniformly distributed in regions free from A-, B- or D-defects.

Recent experimental results on oxygen doped FZ grown Si crystals led Chikawa (1987) to the conclusion that precipitation of residual oxygen is responsible for the formation of A-, B- and I-defects. All these microdefects are formed by a nucleation process, where the nucleation centers are microscopic oxygen clusters. For A- and B-defects the nucleation is heterogeneous. Clusters of intrinsic point defects grow around a nucleus of oxygen atoms. For I-defects the nucleation is homogeneous; both the nucleus and the growing cluster consist of oxygen atoms. Abe et al. (1983) and Chikawa (1987) attempted to analyze the origin of the microscopic oxygen clusters. The unique conclusion from these experimental results is that the formation of dislocation loops originates at the growing melt/crystal interface due to a nucleation process, either homogeneous or heterogeneous, where the nucleation centers are microscopic clusters of atomic size native lattice or impurity defects.

Recently, Tomizawa et al. (1987) suggested that growing GaAs crystals by the CZ method under a controlled As vapor pressure, following a technique introduced by Leung and Allred (1973) leads to significant reduction of the dislocation density. Their interpretation of these experiments was based on the combined effects of the improvement of the melt composition and the reduction of thermal stresses without Ga droplet formation. The improvement of the melt stoichiometry leads to a lower point defect concentration at the melt/solid interface. Analysis of the composition of LEC grown GaAs crystals by X-ray techniques (Fujimoto, 1987) indicated that the atomic concentration of As was higher than that of Ga. As-interstitials at tetrahedral sites surrounded by Ga atoms are the main defects related with nonstoichiometry; these are the EL2 defects mentioned in Section 3.1.

Sumino (1987b) reviewed the effects of impurity doping on dislocation densities in GaAs crystals. There are two main effects other than the retardation of dislocation motion
due to solid solution lattice hardening: locking of the newly generated dislocation loops by impurity precipitation before the loops obtain a critical size and become mobile and formation of impurity-point defect complexes, which inhibit loop formation. These effects are the basis for the production of dislocation free III-V compound semiconductors due to sufficient impurity doping (Sumino, 1987b).

2.5. Dislocation Modeling in Crystal Growth

Dislocation modeling in crystal growth has been traditionally limited to capturing only the qualitative features of the dislocation field in the growing crystal based on the analysis of thermal elastic stresses developed in the crystal due to the nonuniformities of the temperature field. This direct relation to the details of the thermal stress field in the crystal motivated systematic stress analyses and extensive thermal modeling of crystal growth systems. However, according to the analysis of Section 2.4 thermal stress is not the driving force of dislocation generation unless there is already some population of dislocations in the crystal. Quantitative modeling of dislocation dynamics during growth requires simultaneous analysis of elastoplastic deformation and dislocation multiplication and is the subject of Chapter 3.

2.5.1. The CRSS Model for Dislocation Formation

The CRSS method was introduced by Jordan and coworkers (1980) for the calculation of the radial variation of the dislocation density of GaAs crystals grown in the (001) direction. The method was motivated by the early works of Billig (1955) and Penning (1958). Billig concluded from his experiments with Ge crystals grown by the CZ technique, that high thermal stresses gave rise to slipping and dislocation generation and he made order-of-magnitude calculations of thermal stresses and dislocation densities. Penning reported etch pit densities in Ge, showing dislocation densities to be highest at the periphery of the crystal, intermediate in the center and lowest along an internal annular region, and interpreted his results in terms of thermal stresses. He suggested two alternative paths of stress relief in the growing crystal: the complete plastic stress relief due to the generation
and motion of dislocations and the incomplete plastic stress relief, where the thermal stress is mostly elastic and only a small fraction of the thermal stress is released by plastic flow. Jordan et al. (1980) adopted the latter interpretation in their comparison of observed etch pit patterns to the stress in the crystal in excess of the CRSS.

The CRSS method consists of three parts. First, the temperature profile in the growing crystal is calculated using realistic thermal boundary conditions. Second, thermal stresses in the crystal are calculated applying linear thermal elasticity. Finally, the resolved shear stresses in all the slip systems are calculated following tensor transformations given by products of rotation dyadics. For diamond structure crystals, the dislocation density \( N \) is taken to be proportional to the sum of the 12 excess resolved shear stresses \( \{\sigma_i^e\} \) acting in the \( \{111\}\langle1\bar{1}0\rangle \) slip systems:

\[
N \propto \sum_{i=1}^{12} \sigma_i^e.
\]  

(2.12)

The excess shear stress is defined as the difference between the absolute value of the resolved shear stress \( \{\sigma_i\} \) and the CRSS as

\[
\sigma_i^e = \begin{cases} 
|\sigma_i| - \text{CRSS}, & \text{if } |\sigma_i| > \text{CRSS}, \\
0, & \text{if } |\sigma_i| \leq \text{CRSS}.
\end{cases}
\]  

(2.13)

This CRSS model (Jordan, 1979; Jordan et al., 1980, 1981; Jordan and Parsey, 1988) gave good qualitative results for the radial variation of the dislocation densities. A comparison between an experimentally observed dislocation pattern in a GaAs wafer grown by the LEC method, and contours of excess CRSS predicted by the CRSS method is presented in Figure 2.12. However, it is a weak approach for quantitative modelling, because it is based on a continuum linear elastic treatment of a solid with a perfect lattice that is deformed by the action of mechanical stress, while the micromechanical effects of the dislocation motion are not taken into account. The weakness of the method becomes clear, e.g. in the LEC growth of GaAs, where the sum of the right side of (2.12) decreases significantly for axial locations above the encapsulant height (Motakaf and Witt, 1987), while the etch pit densities do not indicate such a decrease in the dislocation density (Jacob et al., 1982). This example shows that rate phenomena are important. Dislocation dynamics has irreversible aspects that cannot be captured by a simple algebraic proportionality such as (2.12).
Figure 2.12. Comparison between (a) an experimentally observed dislocation pattern in a GaAs wafer grown by the LEC method and (b) contours of excess CRSS predicted by simulation.
2.5.2. Advances in Thermal Stress Modeling for CZ and LEC Growth

Although the CRSS model is not adequate for quantitative dislocation modeling, the good qualitative coupling between the role of induced thermal stresses and dislocation formation motivated extensive studies for the improvement of the thermal stress modeling in semiconductor crystals grown by the CZ and LEC techniques. The original analysis of Jordan et al. (1980) was based on a simplified two-dimensional, quasi-steady state model for the temperature field in the grown crystal that included parameters such as the pull rate, convective and radiative heat transfer between the crystal and a constant temperature ambient and thermophysical properties. The calculation of the thermal elastic stress was based on the plain strain assumption (Boley and Weiner, 1960) and hence was approximate for a cylindrical crystal.

Kobayashi and Iwaki (1985) used Love's stress functions to investigate the effects of the Biot ($Bi$) and Peclet ($Pe$) numbers and the crystal length in determining the induced thermal stresses in CZ grown Si. They also showed the inadequacy of the plain strain assumption for stress calculations near the melt/crystal interface. Duseaux (1983) calculated the temperature and thermoelastic stress fields in LEC grown GaAs crystals using the finite element method and studied the effects of the neck region. He used an axial ambient temperature profile with constant temperature gradients in the encapsulant and the inert gas based on experimental data. Szabo (1985) carried out stress calculations in a semi-infinite cylindrical crystal and established the significance of the axial gradient and curvature of the temperature field in the determination of stresses away from the melt/crystal interface. Lambropoulos (1987a) used the finite element method to show that for CZ grown crystals with lengths higher than 3-4 radii all stress components achieve steady state values, and that all significant stress variation occurs within one radius from the melt-solid interface. Motakef and Witt (1987) established the significance of the encapsulant geometry and transparency on the resulting thermal stresses in GaAs crystals. All the authors cited above used traction free boundary conditions for the thermal stress calculations and assumed that the growing crystal is an isotropic solid. Lambropoulos (1987b) determined the extent of the assumption that a CZ grown Si crystal is elastically isotropic.
He compared results of thermal stress calculations using the finite element method both under the isotropic assumption and by fully accounting for the elastic anisotropy of crystals grown along the (100) and (111) directions. He concluded that the isotropic assumption underestimates the thermal stresses by as much as 20%.

Calculations of the temperature and the thermal stress fields in a growing crystal are necessary for quantitative defect modelling. Numerical simulation using realistic thermal boundary conditions provides reliable information for the details of the temperature and thermal stress profiles and indicates strategies for the reduction of defect densities. The large scale computational work of Thomas (1988) for batchwise simulation of the LEC growth of GaAs, including convective and radiative heat transfer and free interface shapes, is an excellent example for the amount of information that can be obtained by numerical simulation compared to simplified approximate calculations. This information is invaluable for the systematic study of the effects of processing conditions on defect densities in the material (Thomas, 1988; Bornside et al., 1991; Kinney, 1992).

2.5.3. Modeling of Inelastic Effects

Lambropoulos et al. (1983) first attempted more realistic modeling of crystalline stress than the usual thermal elasticity approach to calculate stresses that arise in Si ribbon growth techniques, where Si sheets grow at high speed and under high temperature gradients. To account for the effects of plasticity, these authors used a constitutive equation for the elastic-creeping behavior of Si. This equation was a modification of the model of Myshlyaev et al. (1969), who studied the creeping behavior of Si under uniaxial compression at high loads $\sigma$ and temperatures in the range of 900 to 1300°C. According to this model, the steady creep rate of Si crystals is governed by the kinetic equation

$$\dot{\varepsilon} = \dot{\varepsilon}_0 \exp\left(-\frac{U_0 - V\sigma}{kT}\right),$$

(2.14)

where $U_0$ is an energy barrier higher than the Peierls barrier, $V$ is an activation volume and $U_0$ is related to the sublimation energy and the activation energy for vacancy diffusion.

Another modification of the same model was used by Lambropoulos (1987a) to study the effect of plastic deformation in the CZ growth of Si. In both cases (Lambropoulos
1987a, 1983), the calculated stresses were lower than the ones calculated if the crystal was
considered to be elastic, with a different stress distribution near the melt/crystal interface.
In spite of the more realistic predictions, this model of creeping behavior is not sufficient for
quantitative dislocation modelling, because the strain rate is not related to the dislocation
density as in Orowan's equation, eq. (2.4).

A totally different approach for the modeling of plastic deformation dynamics has been
followed by Alexander and Haasen (1968). This approach is centered on the coupling of
plastic deformation with dislocation propagation and multiplication including dislocation
interactions that give rise to internal stresses in the crystal. The model of Alexander and
Haasen (1968) emphasizes the importance of the incorporating the dislocation density as a
variable in the analysis of plastic deformation. The model successfully predicts the results
of static creep tests performed with constant applied stress and temperature, and dynamic
deformation experiments performed with constant strain rate and temperature. According
to the model, the upper and lower yield stresses of the crystal are given by

\[ \tau_y = C_y \varepsilon^{1/n} \exp\left(\frac{U}{kT}\right), \]  

(2.15)

where \( C_y \) is a constant different for the upper and lower yield point respectively, \( n = m + 2 \)
and \( U = Q/n \), where \( Q \) is the activation energy for dislocation glide, given in eq. (2.5).
According to equation (2.15) and the discussion of section 2.1.2, the CRSS is a function of
both temperature and strain rate, this dependence is not taken into account by the CRSS
model. The model of Alexander and Haasen (1968) is presented in detail in Section 3.

73
3. QUANTITATIVE DISLOCATION MODELING IN CRYSTAL GROWTH

The objective of quantitative dislocation modeling in bulk melt growth systems is the accurate prediction of the dislocation density distribution in the growing crystal. Calculation of the density of grown-in dislocations provides the initial conditions for modeling dislocation motion and multiplication during the subsequent stages of device processing. In addition, quantitative analysis of the effects of processing conditions on the dislocation density of a single crystal produced by bulk melt growth provides the thermomechanical criteria that determine processing strategies for the reduction of the dislocation density in the crystal. However, modeling of dislocation dynamics during growth is not an easy task. Understanding of the microscopic aspects of the dislocation core structure and dynamics is required for the development of macroscopic models that describes the deformation dynamics of the growing crystal.

The details of the structure and dynamics of dislocations in the diamond/zincblende crystallographic structure are analyzed in Section 3.1 and a general macroscopic equation is developed for the calculation of the dislocation density in the bulk crystal. A model of dislocation dynamics during semiconductor crystal growth is developed in Section 3.2 based on the general macroscopic equation of Section 3.1, on ideal plasticity theory and on the constitutive equations of Alexander and Haasen (1968) for the dynamics of plastic deformation caused by mobile dislocations in the diamond/zincblende structure. The appropriate scaling analysis for the dislocation density and the plastic strain is carried out in Section 3.3 and the dislocation density is computed in the asymptotic limit of rapid dislocation multiplication, which is the case for typical CZ and LEC growth conditions. The implications of the asymptotic analysis for the CZ growth of silicon and the LEC growth of GaAs and InP are presented in Section 3.4.

3.1. Dislocation Structure and Dynamics in Diamond/Zinblende Structure Crystals

3.1.1. Microscopic Aspects

The mobile dislocation configuration in the diamond/zincblende structure was pre-
sented in Section 2.2.1 as revealed by chemical etching and X-ray topography (Sumino and Harada, 1981). More recent experimental and theoretical studies have clarified the structural details of the dislocations of the mobile configuration. These studies have been reviewed by Alexander (1986). The foundation of quantitative dislocation modeling is provided by eqs. (2.4), (2.5) and (2.7). These equations are supported by microscopic experimental and theoretical evidence. The recent advances in the understanding of the core structure and dynamics are reviewed below.

The diamond structure is a lattice with a two-point basis consisting of two fcc sublattices (Ashcroft and Mermin, 1976). This leads to a layered structure, each layer consisting of two (111) planes connected by three covalent bonds per atom, as shown in Fig. 3.1. The number of bonds between two planes belonging to neighboring layers is one per atom and the distance between two layers is three times as large as the distance within a layer. Our attention is focused on the two principle types of dislocations in the strong Peierls potential of the covalent crystal, the minima of which run parallel to the (110) directions. These are the 60° dislocation and the screw dislocation. These types of dislocations lie along (110) directions and control the early stages of crystallographic glide (Hirth and Lothe, 1982; Alexander, 1986). Our interest is concentrated on the initial stages of deformation ($\varepsilon \leq 5\%$) caused by the thermomechanical conditions of processing.

Shear displacement, accompanied by dislocation motion, may operate between two layers or between the two (111) planes constituting one layer. The two types of glide planes are labeled as I and II in Fig. 3.1. Dislocations moving on glide planes of type I are called shuffle dislocations, while those moving on type II planes are glide dislocations (Hirth and Lothe, 1982; Alexander, 1986). Complete dislocations such as the 60° and screw dislocations shown in Fig. 2.6, tend to dissociate into partial dislocations with Burgers vectors of smaller magnitudes to minimize the energy of the crystal by relaxing the strain field of the complete dislocation (Hirth and Lothe, 1982). The glide dislocation dissociates into two Shockley partial dislocations in exactly the same manner as in fcc metals (Alexander, 1986). The dissociation is described symbolically as

$$\frac{a}{2}[\overline{1}01] \rightarrow \frac{a}{6}[\overline{2}11] + \frac{a}{6}[\overline{1}12]. \quad (3.1)$$
Figure 3.1. (a) View of a diamond lattice with the (111) planes perpendicular to the plane of the paper. (b) Projection of the diamond lattice onto the (110) plane. The two fcc sublattices and the two types of glide planes (I and II) are shown. (c) Projection of the diamond lattice onto the (111) plane. The Burgers vector $\mathbf{b}$ of the complete dislocation and the Burgers vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ of the two partial dislocations are shown (Alexander, 1986).
The Burgers vector $\mathbf{b}$ of the complete dislocation and the Burgers vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ of the corresponding partial dislocations are shown in Fig. 3.1. The dissociation of a shuffle dislocation described by Alexander (1986) is a more complicated process. The 60° dislocations dissociate into dipoles of 30° and 90° dislocations, while screw dislocations dissociate into dipoles of two 30° partial dislocations.

Glide and shuffle dislocations differ in the directions of broken bonds occurring in the dislocation cores. Alexander (1986) described how the presence of any shuffle dislocation in unrelaxed configurations leads to a row of dangling bonds, which are directed perpendicular to the glide plane; however, any glide dislocation contains one or more lines of broken bonds which are parallel to the dislocation line or weakly inclined to the glide plane. The perfect screw dislocation of the shuffle set may be constructed without any dangling bonds and the dissociated screw, consisting of two 30° partials, contains only dangling bonds parallel to the dislocation line (Alexander, 1986). These configurations of dislocation cores have important implications in the electrical activity of dislocations and their interactions with electrically active impurities and point defects.

High resolution transmission electron microscopy (TEM) clearly shows nearly all glissile dislocations in elemental semiconductors and in III-V compounds to dissociate into two partial dislocations bounding a stacking fault ribbon (Alexander, 1986). Moreover, imaging dislocations moving in the electron microscope proved that the dislocations are mobile in the dissociated state. Apparently the larger number of bonds and the smaller distance between the lattice planes within one layer of (111) planes do not preclude type II glide planes from operating.

For each type of dislocations, e.g. 60° dislocations or screw dislocations, and for a specific glide system the dislocation velocity is very well described by eq. (2.5)

$$v = B_0 \exp\left(-\frac{Q}{kT}\right) \tau^m,$$

where $B_0$ is a material constant, $Q$ is constant in the stress range of interest, typically of order 10 MPa, and $m$ is slightly temperature dependent (Alexander, 1986). In spite of both the experimental and the theoretical progress in the field during the past ten years there is no satisfactory theory for the prediction of the mobility coefficient and
the stress exponent $m$. Therefore, the parameters $B_0$ and $m$ are treated as empirical phenomenological constants. On the contrary, the activation energy $Q$ is satisfactorily described by the theory of double kink formation presented in Section 2.2.2.

In silicon, both 60° and screw dislocations have equal velocities (Alexander, 1986; Sumino, 1987a, 1987b). This is not the case in III-V compound semiconductors (Sumino, 1987b) where three dislocation types with different velocities must be distinguished: $\alpha$ and $\beta$ 60° dislocations and screw dislocations. The two arms of 60° dislocations in the mobile configuration of Fig. 2.6 are of $\alpha$ and $\beta$ type, respectively. Therefore, 60° dislocations of plus or minus sign depending on their orientation with respect to the Burgers vector are characterized as $\alpha$ and $\beta$, respectively. A 60° dislocation has a row of atoms of the Group-III element or the Group-V element at its core, i.e. at the edge of the extra half atomic plane it accompanies depending on the sign of the Burgers vector. The dislocations with atoms of the Group-III element at the core are the $\alpha$ dislocations and those with atoms of the Group-V element at the core are the $\beta$ dislocations (Sumino, 1987b). The velocity of $\alpha$ dislocations has been shown to be higher than the velocity of $\beta$ dislocations by as much as two orders-of-magnitude over the temperature range 673 K $< T <$ 823 K, while screw dislocations have intermediate velocities between the two types of 60° dislocations (Sumino, 1987b).

3.1.2. Atomistic Simulations of Dislocation Structure and Dynamics

The structure, energetics and dynamics of dislocations in metals and covalent crystals have been studied by atomistic simulations following the methods of lattice statics and molecular dynamics (Duesbery, 1989). Atomistic studies of dislocation dynamics using molecular dynamics techniques have been limited to times $O(1 \text{ ns})$. The embedded-atom interatomic potential (Daw and Baskes, 1984) has been used for the study of dislocation dynamics in fcc metals (Baskes et al., 1988, 1989). The calculations were carried out for low temperatures and very high stresses. Under such conditions dislocations move at very high speeds that approach the speed of sound in the metal. The time scale is set by the dislocation velocity and the linear dimension of the simulation cell limited to $O(10)$ atomic
spacings. The molecular dynamics results (Baskes et al., 1988, 1989) were in excellent agreement with the continuum model for relativistic speed dislocation dynamics (Hirth and Lothe, 1982).

Nonequilibrium molecular dynamics techniques have been used to study high strain rate plastic flow in model solids (Ladd and Hoover, 1983; Ladd, 1985). The limitation to times $O(1 \text{ ns})$ makes atomistic simulations unfeasible for plastic strain rates less than $10^{10} \text{ s}^{-1}$, an incredibly high value. For constant strain rate simulations the time scale is set by the inverse of the strain rate. The introduction of mesoscopic dynamics made practical simulations with plastic strain rates $O(10^8 \text{ s}^{-1})$ (Ladd, 1985; Hoover, 1986). Mesoscopic dynamics are intermediate simulations in which the dislocations are the fundamental particles replacing the underlying atomic coordinates and momenta. Instead of using an interatomic potential to determine the positions and momenta of atoms from Newton's equations of motion for the atoms, equations for the interactions between dislocations are used to determine the positions and velocities of simple dislocation configurations. Thus, the number of degrees of freedom decreases drastically making feasible computations for times higher by two orders-of-magnitude (Ladd, 1985; Hoover, 1986).

Atomistic simulation of dislocation dynamics in the diamond structure is necessary to clarify the microscopic mechanisms of dislocation propagation. Unfortunately, atomic modeling of dislocations in the diamond structure is limited by the time scales associated with dislocation dynamics under typical deformation conditions. Deformation of diamond structure crystals during crystal growth and device processing is characterized by strain rates $O(10^{-4} \text{ s}^{-1})$ (Alexander and Haasen, 1968). Such strain rates correspond to time scales on the order of $10^4 \text{ s}$ which makes molecular dynamics simulations of dislocation propagation totally impractical. Molecular dynamics simulation becomes an even more impractical task during growth under CZ or LEC conditions where the growth rates are extremely small $O(1 \text{ cm/h})$. Molecular dynamics simulations of propagation of the melt/solid interface of silicon have been limited to growth velocities $O(10^2 \text{ m/s})$ which are characteristic of laser melting (Landman et al., 1986; Broughton and Abraham, 1986; Grabow et al., 1989). However, targeted lattice statics simulations for the structure and energetics of dislocation core are feasible and shed new light to dislocation dynamics provided that
accurate interatomic potentials are available (Duesbery, 1989).

Atomistic studies of dislocation cores in silicon were presented recently by Duesbery et al. (1991), who used lattice statics simulations. Analysis of dislocation mobility requires simulation cells containing thousands of atoms because of the long-range strain field created in the lattice by the dislocation. Duesbery et al. (1991) used simulation cells that contained 720 to 3780 atoms and the many-body interatomic potentials of Stillinger and Weber (1985), Tersoff (1989), and Kaxiras and Pandey (1988) in their lattice statics simulations. Both 60° and screw dislocations were considered and modeled as partial dislocation dipoles with a separation of 2α, where α is the lattice parameter. In all cases the configurations were equilibrated by constant pressure molecular dynamics at T = 100 K and subsequently quenched to 1 K to extract ground-state energies.

Calculations with all three interatomic potentials predicted reconstructed core structures with four-fold coordination, which is possible by symmetry-breaking deformation (Duesbery et al., 1991) and is in agreement with experiment (Alexander, 1986) and band structure calculations (Heggie and Jones, 1983a, b; Chelikowsky and Spence, 1984). Calculation of the formation energy of a kink pair is complicated by the fact that for the 30° and 90° partial dislocations kink pairs may be fully reconstructed or may be combined with an antiphase defect or soliton in the broken-symmetry core. The reported formation energies pertain to kink pairs containing a single dangling bond because the motion of reconstructed kinks requires the intermediate creation of dangling bonds. However, all three potentials predict formation energies of kink pairs larger for the 30° partial than the 90° partial by a factor of 1.5 to 4 depending on the interatomic potential. This would imply a much lower mobility according to eq. (2.5) for screw dislocations which contain two 30° partials than for 60° dislocations which contain one 30° and one 90° partial; this prediction does not agree with experimental data (Alexander, 1986; Sumino, 1987a) as reviewed in Section 3.1.1. Therefore, none of the empirical interatomic potentials is completely satisfactory for description of dislocation mobility. These results emphasize the need to include the wide range of stable, highly deformed configurations in the dislocation core in the database for fitting empirical interatomic potentials (Duesbery et al., 1991).

All the attempts at atomic modeling motivate the formulation of a rigorous macro-
scopic model of dislocation dynamics taking into account the well-established microscopic aspects of dislocation structure and dynamics. This is particularly necessary for the study of dislocation dynamics during bulk melt growth.

3.1.3. Macroscopic Model of Dislocation Dynamics

The first attempt to quantitatively model the dynamics of grown-in defects was presented by Keizer et al. (1985) for the growth of ice crystals. The model was motivated by data from anomalous light scattering experiments (Güttinger et al., 1979) and for the shrinkage of faulted and unfauluted dislocation loops (Hondoh et al., 1983). Experimental results were available due to the optical transparency of the material. This work provides a useful framework for the study of dislocation dynamics via a macroscopic transport equation for the number density of mobile dislocations per unit volume.

The dynamics of mobile dislocations is described by a transport equation for the number density $N_i$ of each mobile dislocation species $i$ per unit area. The experimental observation of equal velocities of the different segments of the mobile dislocation configuration in silicon (Alexander and Haasen, 1968) permits the description of dislocation dynamics by a single variable $N$ that corresponds to the number density of the entire mobile dislocation. Although multiple dislocation species are needed to describe dislocation dynamics in III-V compound semiconductors, a single variable $N$ is retained in the analysis to simplify the notation. All macroscopic dislocations are considered mobile in this treatment as a result of the ability of dislocations to move both at the undissociated and the dissociated state (Alexander, 1986).

The macroscopic transport equation for the number density $N$ of mobile dislocations per unit area is written as

$$\frac{\partial N}{\partial t} = -\nabla \cdot C_N - \nabla \cdot D_N + G_N + M_N - L_N,$$

where $C_N$ is the dislocation convective flux, $D_N$ is the dislocation diffusive flux, $G_N$ is the rate of dislocation generation per unit area due to dislocation sources in the bulk crystal, $M_N$ is the rate of dislocation multiplication per unit area and $L_N$ is the rate of dislocation
destruction per unit area. The rate of dislocation destruction $L_N$ is decomposed as

$$L_N = L_{1,N} + L_{2,N},$$

(3.3)

where $L_{1,N}$ is the rate of destruction of dislocation loops due to shrinking and $L_{2,N}$ is the rate of dislocation destruction due to pair annihilation.

The rate of dislocation generation is set equal to zero ($G_N = 0$). Generation may be due to external factors, such as crystal surface damage (Dash, 1959). However, such factors are controlled during processing and are not taken into account in the analysis. According to the discussion in Section 2.4, dislocation generation from sources other than mobile dislocations is limited in the melt/crystal interfacial region where dislocations are nucleated from clusters of supersaturated point defects. The main source of dislocation formation in the bulk is dislocation multiplication expressed by $M_N$.

The convective flux of dislocations $C_N$ in equation (3.2) is formulated in a frame of reference where the crystal is fixed, i.e., a frame of reference that moves with velocity $V_{f.o.r.} = -V_g$,

$$C_N \equiv (v - V_{f.o.r.})N = (v + V_g)N,$$

(3.4)

where $v$ is the dislocation velocity and $V_g$ is the crystal growth rate. The growth velocity $V_g$ is taken equal to the crystal pull rate $V_p$, which sets the time scale for crystal translation. The divergence of the convective flux for a continuous distribution of dislocations in a fixed frame of reference is equal to zero

$$\nabla \cdot vN = 0.$$  

(3.5)

The explanation for the result of equation (3.5) is trivial if one considers only straight edge dislocation lines and a control volume composed of either a disc with infinitesimal thickness and surfaces parallel to the glide plane or a plane perpendicular to the glide plane and parallel to the dislocation lines. Since the dislocation lines do not intersect the control volume, there is no change in the dislocation density per unit area inside the control volume due to the dislocation glide motion.

This very simple geometrical proof has been generalized for a random continuous distribution of dislocations (Nabarro, 1987). In the notation used here eq. (3.5) is equivalent
to the general equation of dislocation field theory

\[ \nabla \cdot \mathbf{a} = 0, \]

(3.6)

where \( \mathbf{a} \) is Kröner's dislocation density tensor (Kröner, 1962) and \([a]_{jk}\) is the sum of the \( b_j \) components of the dislocations with Burgers vector \( \mathbf{b} \) threading unit area normal to \( k \). Equation (3.6) states that dislocations do not end inside the crystalline medium. It is easily proved, because

\[ \mathbf{a} = -\nabla \cdot \mathbf{N}, \]

(3.7)

where \( \mathbf{N} \) is the dislocation movement triadic defined as

\[ \mathbf{N} = \mathbf{l} \mathbf{b}, \]

(3.8)

and \([l]_{ij}\) is the projection in the \( j \)-direction of the length of a dislocation line, which crosses an area normal to the \( i \)-direction (Nabarro, 1987). This notation is shown in Fig. 3.2.

If complete dislocation lines in the diamond structure were diffusing, a surface diffusive flux would be defined by the equation

\[ \mathbf{D}_N = -\mathbf{D}_d \cdot \nabla \mathbf{N}, \]

(3.9)

where \( \mathbf{D}_d \) is the dislocation diffusivity tensor. Keizer et al. (1985) defined a diffusive flux \(-D \nabla n_l\) for microscopic faulted loops in ice, where \( n_l \) is the volume number density of the loops. The importance of this transport mechanism was inferred from light scattering measurements (Güttlinger et al., 1979) that determined the diffusion of particles in crystalline ice (Güttlinger et al., 1979). Keizer et al. assumed that the light scatterers were microscopic dislocation loops in the ice.

Mobile macroscopic dislocation loops in the diamond structure do not diffuse by a stochastic process (Alexander and Haasen, 1968). Therefore, the diffusive flux of equation (3.9) is set equal to zero. It should be mentioned that diffusive mechanisms of kink dynamics, which determine the magnitude of the dislocation velocity (see Section 2.2.2), are not included in this analysis. The macroscopic equation (3.2) governs the evolution of the dislocation density and not the density of microscopic steps on dislocation lines.
Figure 3.2. Geometrical interpretation of the notation used in the treatment of random continuous distributions of dislocations.
The kinetics of dislocation multiplication is governed by equation (2.7) and is analyzed in detail in Section 3.2.1. Keizer et al. (1985) attributed the destruction of dislocation loops to shrinking caused by the emission of native lattice point defects from the loops, according to the model

$$L_\phi = \tau^{-1}n_l,$$

(3.10)

where $\tau$ was the annealing lifetime of a dislocation loop and $n_l$ is the density of dislocation loops. The first order kinetics expressed by eq. (3.10) describes the emission of point defects from the loop according to the approximate analysis of Birnbaum et al. (1971). On the other hand, Hondoh et al. (1983) interpreted the results of their experiments using the model of Seidman and Balluffi (1966) for the annealing by climb of a prismatic dislocation loop in the center of a thin foil. Such a loop shrinks under the influence of its line tension by emitting vacancies which diffuse to the surface of the thin foil. However, loop shrinking is an important effect for unstabilized loops, i.e. microscopic faulted loops with radius less than a critical radius (Alexander and Haasen, 1968). These loops are not mobile in the diamond structure. The mobile dislocation loops in the diamond structure are stabilized so that the line tension in the loop is always compensated by the applied stress in the crystal (Schäfer, 1967; Alexander and Haasen, 1968). For this reason, the shrinkage of mobile dislocation loops in the diamond structure is not taken into account and the component $L_{1,N}$ in equation (3.3) vanishes, i.e.

$$L_{1,N} = 0.$$

(3.11)

Unstabilized loops, which shrink by emission of vacancies, exist in the crystal very near the melt/solid interface, where loop shrinkage is an opposing force to mobile dislocation generation.

Dislocation annihilation occurs at the junction of two loops with opposite Burgers vectors or two portions of the same loop. Johnston and Gilman (1959) attempted to explain why the dislocation density reaches a final value as time increases and subtracted a term proportional to $N^2$ from the right hand side of equation (2.7) to account for annihilation. Second order kinetics of dislocation annihilation corresponds to the interaction of dipoles of complete dislocations instead of single dislocation lines (Alexander and Haasen, 1968)
and are not accounted for in this model. The details of dislocation annihilation kinetics are discussed in Section 3.2.1.

Combining equations (3.2)-(3.5), (3.9) and (3.11) gives

\[ \frac{\partial N}{\partial t} = -\nabla \cdot \mathbf{V}_p N + M_N - L_{2.N} \]  

(3.12)

Equation (3.12), coupled with the equations for the thermal elastoplastic stress field, is the basis for the analysis of dislocation dynamics for the LEC growth of GaAs and InP and the CZ growth of Si. The absence of diffusive terms makes (3.12) a first order differential equation and only one boundary condition next to the melt/crystal interface \( z = 0 \) is necessary for its solution. The growing crystal is decomposed into two regions, as it is shown in Figure 3.3. The region \( z \geq 0 \) corresponds to the bulk crystal, where dislocation dynamics is governed by equation (3.12). The region \(-d \leq z < 0\) corresponds to an interfacial region of thickness \( d \) where dislocations are generated. Modeling of the kinetics of dislocation generation in this region provides the appropriate boundary condition, \( N(0) = N_0 \), for the bulk crystal model.

3.2. Modeling of Dislocation Dynamics During Semiconductor Crystal Growth

Our modeling of dislocation dynamics in semiconductor crystals during bulk melt growth is based on the analysis of Section 3.1.3 and the phenomenological model of Alexander and Haasen (1968) for the dynamics of plastic deformation in the diamond/zincblende structure. The model of Alexander and Haasen (1968) describes the motion and multiplication of dislocations in the presence of an applied stress field and relates the formation and interaction of dislocations to the rate of generation of plastic strain in the material. The scalar form of the model for the study of uniaxial deformation experiments is presented in Section 3.2.1., while a three-dimensional formulation is developed in Section 3.2.3. The model of Alexander and Haasen (1968) has been successfully used to predict the results of static creep experiments performed under constant stress and temperature, and dynamic deformation experiments performed under constant strain rate and temperature (Alexander and Haasen, 1968; Haasen, 1967; Peissker et al., 1962; Schröter et al., 1983;
Figure 3.3. Decomposition of the growing crystal into an interfacial region and a bulk region where dislocation dynamics is analyzed.
Müller et al., 1985). Although much attention has gone into understanding plastic deformation in crystal growth using the model of Alexander and Haasen, until recently, the calculations of Dillon et al. (1986, 1987) for analysis of thermal stresses in high-speed silicon ribbon growth were the only application of this dislocation dynamics model to a crystal growth system.

Recently, Völkl and Müller (1989) have demonstrated the success of the model of Alexander and Haasen for predicting dislocation profiles in III-V compound semiconductors by predicting the final dislocation density in InP crystals grown by the LEC method. The results of the numerical calculations of Völkl and Müller (1989) for the dislocation density distribution in the crystal for differing growth conditions were in quantitative agreement with etch-pit profiles from crystals grown using conditions that corresponded to the input temperature field in the crystal. Thomas (1988) carried out similar calculations for analysis of dislocation formation in GaAs crystal growth.

The picture of dislocation formation that emerges from analysis of the model of Alexander and Haasen differs from the approach introduced by Jordan et al. (1980) based on the CRSS model. Although the qualitative aspects of basing the dislocation density on the stress in excess of the CRSS are reasonable, the complexity of the microscopic dynamics of dislocations, as described in Sections 2.2 and 3.1, is not completely described by this approach. The CRSS model neglects all the important micromechanical rate processes and is limited to qualitative descriptions of the dislocation field in the crystal as pointed out in Section 2.5. Additional comparison between the model of Alexander and Haasen and the CRSS model is carried out in Section 3.4.

3.2.1. The Model of Alexander and Haasen

The model of Alexander and Haasen (1968) is a set of phenomenological constitutive equations that couple a kinetic rate expression for the formation of dislocations with a model for plastic stress relief in the crystal accompanied by the internal stresses that are created by the dislocation field. The Haasen model describes the evolution of the plastic component of the strain tensor $\varepsilon''(x,t)$ in terms of the dislocation density field $N(x,t)$,
where the position vector $\mathbf{x}$ denotes a point in space. The model is first described in its scalar form, as it was formulated for describing unidirectional deformation experiments (A. E. E. and Haasen, 1968; Haasen, 1967). The model is generalized to a tensorial analog in Section 3.2.3.

The rate-of-change of the plastic strain is related to the density of mobile dislocations by the well-known Orowan equation (Orowan, 1940), i.e. equation (2.4), which is written here as

$$\frac{d\epsilon^p}{dt} = b\dot{v}N, \quad (3.13)$$

where $\epsilon^p$ is the component of the plastic strain tensor resolved in the slip plane and along the slip direction during uniaxial deformation, $\dot{v}$ is the average dislocation speed in the slip system, and $b$ is the magnitude of the Burgers vector of the dislocations. Equation (3.13) is different from equation (2.4) in that the dislocation density is used instead of the mobile dislocation density. This is explained by the fact that all macroscopic dislocations are assumed to be mobile in the diamond/zincblende structure, as discussed in Section 3.1. Immobile dislocations are unstable and the evolution of their density is not governed by the general macroscopic equation (3.12). In the case of silicon, equation (3.13) is rigorously valid for the resolved plastic strain over a particular slip system due to the equal velocities of $60^\circ$ and screw dislocations. In III-$\overline{V}$ compound semiconductors the difference in the velocities among different types of dislocations implies that the factor $\dot{v}N$ should be substituted by the sum of factors $\dot{v}_iN_i$, where $i$ denotes the dislocation type: $\alpha$, $\beta$, or screw. Yonenaga et al. (1987) proved that equation (3.13) is appropriate for analysis of these crystals, where $N$ is the total dislocation density, i.e. the sums of the densities of the three types of dislocations and $\dot{v}$ corresponds to the velocity of the screw dislocations. This was proved (Yonenaga et al., 1987) using the geometry of the mobile dislocation loop shown in Fig. 2.6 and the ordering of the energies of the different dislocation types. In all cases the magnitude $b$ corresponds to the Burgers vector of the complete dislocation loop in Fig. 2.6.

The dislocation speed is modeled by equation (2.5). Again a single dislocation speed is used independent of the velocities of various dislocation types. Equation (2.5) is rewritten
\[ v = B_0 \exp\left(-\frac{Q}{kT}\right) \tau_{eff}^m. \]  

(3.14)

where the applied stress \( \tau \) of equation (2.5), i.e. the component of the applied stress tensor resolved in the slip plane and along the slip direction, has been replaced by the effective stress \( \tau_{eff} \).

The effective stress \( \tau_{eff} \) that appears in eq. (3.14) is the difference between the magnitude of the applied stress \( \tau \), and the internal or 'back stress' which is developed due to the interactions between dislocations and that causes lattice hardening of the crystal (Haasen, 1962). Haasen (1962) introduced the phenomenological expression that the back stress is proportional to \( N^{1/2} \). This is motivated by the long-range perfectly elastic interaction between two straight dislocations which falls as \( 1/r \), where \( r \) is the distance between the two dislocation lines. For a regular array of dislocations characterized by a density \( N \) the inverse distance \( 1/r \) is equal to the square root of the dislocation density. In particular, the mean internal elastic stress \( \tau_i \) produced by a statistical arrangement of \( N \) parallel edge dislocations is

\[ \tau_i = \frac{Gb}{2\pi(1-\nu)} \sqrt{N}, \]  

(3.15)

where \( G \) is the shear modulus and \( \nu \) the Poisson ratio of the crystal (Hirth and Lothe, 1982; Haasen, 1962). For \( N \) parallel screw dislocations the corresponding internal stress is given by

\[ \tau_i = \frac{Gb}{2\pi} \sqrt{N}, \]  

(3.16)

If the internal stress is higher than the applied stress, \( \tau_{eff} \) is zero making the dislocation speed \( v \) zero, i.e.

\[ \tau_{eff} \equiv \begin{cases} \tau - A\sqrt{N}, & \tau > A\sqrt{N} \\ 0, & \tau \leq A\sqrt{N} \end{cases} \]  

(3.17)

where the constant \( A \) is the strain hardening factor Haasen (1962). Equation (3.16) expresses the dominant mechanism of the lattice resistance to dislocation motion to be the result of the interaction between dislocations and the resulting dislocation retardation according to equation (3.14).
A population of dislocations is modeled by the dislocation multiplication law originally introduced by Johnston and Gilman (1959) and given by equation (2.7). Alexander and Haasen (1968) used the experimental evidence provided by Low and Turkalo (1962) for the initiation of the cross-slip multiplication process and postulated that the kinetic constant $\beta$ of eq. (2.7) is proportional to the effective stress. Therefore, dislocation multiplication kinetics is expressed by

$$\frac{dN}{dt} = K \tau_{eff} vN,$$

(3.18)

where the constant $K$ is determined from deformation experiments under constant load. In addition, equations (3.14), (3.16) and (3.17) guarantee the saturation of the dislocation density in the crystal during deformation. Annihilation of dislocation pairs is modeled directly by using the effective stress $\tau_{eff}$ in the multiplication kinetics. Pair annihilation is driven by the interactions between single dislocation lines which gives rise to the back stress $A\sqrt{\lambda}$ in eq. (3.17). Therefore, second order kinetics, corresponding to the interaction between complete dislocation dipoles, need not be introduced (Alexander and Haasen, 1968).

### 3.2.2. Parameter Estimates

Solution of the rate equations (3.13) and (3.18) of the model of Alexander and Haasen (1968) for the dynamics of plastic deformation in the diamond/zincblende structure with constitutive equations (3.14) and (3.17) requires estimates of the phenomenological constants ($B_0, Q, m, A, K$), the coefficient of thermal expansion and the elastic constants of the crystalline material. We compile here estimates for these parameters for silicon and for the III-V compound semiconductors GaAs and InP.

Yonenaga and Sumino (1978) and Suezawa et al. (1979) presented the results of deformation experiments for silicon over a temperature range of 873-1223 K and reported measurements for all the parameters in the model of Alexander and Haasen. Data is much more scarce measured closer to the 1683 K melting temperature of silicon is much more scarce. Schröter et al. (1983) reported measurements of the activation energy $Q$ and the effective stress exponent $m$ for temperatures up to 1573 K and reviewed previous
experimental results for silicon and germanium. This data, extrapolated to the melting temperature, is listed in Table 3.1. The elastic behavior of silicon is assumed to be isotropic and is expressed by the shear modulus, which is available as a function of temperature from both experimental measurements (Landolt-Börnstein, 1979; Dillon et al., 1987) and molecular dynamics simulations (Klude et al., 1986). The Stillinger-Weber interatomic potential (Stillinger and Weber, 1985) was used in the molecular dynamics simulations of Kluge et al. (1986). The elastic constants were calculated at temperatures $T = 888, 1164$ and $1477$ K. The calculated shear modulus is smaller than the experimental values (Landolt-Börnstein, 1979) by 30% at $T = 888$ K and by 40% at the two higher temperatures. The calculated elastic constants show the same softening with rising temperature as shown by the experimental values (Klude et al., 1986).

Völkl and Müller (1989) reported the parameters for the model of Alexander and Haasen for InP determined from deformation experiments for temperatures up to 0.8 $T_m$, where $T_m = 1335$ K; these measurements are listed in Table 3.1. Deformation and dislocation velocity measurements are lacking for GaAs. Yonenaga et al. (1987) reported values of $Q$ and $m$ measured in the temperature range 673-773 K, far below the 1511 K melting temperature; these results are listed in Table 3.1. These authors do not report measurements of either the dislocation velocity or the multiplication constant $K$. Steinhardt and Haasen (1978) reported velocity measurements for the three different types of dislocations in GaAs at temperatures below 973 K, but did not list values for either the effective stress or the dislocation density, so the constants $B_0$ and $A$ cannot be calculated from their data. Motakef (1991) estimated the constants $A$ and $B_0$ for undoped GaAs based on eq. (3.15), measurements for the lower yield stress, and the complete form of eq. (2.15) of the theory of the initial stage of deformation (Alexander and Haasen, 1968). These values also are listed in Table 3.1. However, Motakef (1991) assumed that the value of the multiplication rate constant $K$ for GaAs is equal to the corresponding value of InP. This conclusion is highly unlikely because the multiplication rate constant is directly related to the mechanical properties of the material (Alexander and Haasen, 1968). If the mechanical properties of InP and GaAs were the same, the model of Alexander and Haasen would suggest that the evolution of plastic strain and dislocation density would be the same. However,
<table>
<thead>
<tr>
<th>Property</th>
<th>Si</th>
<th>GaAs</th>
<th>InP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_m$ (K)</td>
<td>1683</td>
<td>1511</td>
<td>1335</td>
</tr>
<tr>
<td>m</td>
<td>1.1</td>
<td>1.7</td>
<td>1.4</td>
</tr>
<tr>
<td>$Q$ (eV)</td>
<td>2.2</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>b (m)</td>
<td>$3.8 \times 10^{-10}$</td>
<td>$4.0 \times 10^{-10}$</td>
<td>$4.2 \times 10^{-10}$</td>
</tr>
<tr>
<td>A (N m$^{-1}$)</td>
<td>4.0</td>
<td>5.2</td>
<td>3.0</td>
</tr>
<tr>
<td>K (m N$^{-1}$)</td>
<td>$3.1 \times 10^{-4}$</td>
<td>-</td>
<td>$1.2 \times 10^{-2}$ ($T_m$)</td>
</tr>
<tr>
<td>$B_0$ (m$^{2m+1}$ N$^{-m}$ s$^{-1}$)</td>
<td>$8.6 \times 10^{-4}$</td>
<td>$1.8 \times 10^{-8}$</td>
<td>$1.36 \times 10^{-8}$ ($T_m$)</td>
</tr>
<tr>
<td>$G$ (N m$^{-2}$)</td>
<td>$3.51 \times 10^{10}$</td>
<td>$4.0 \times 10^{10}$</td>
<td>$2.36 \times 10^{10}$ (1200 K)</td>
</tr>
</tbody>
</table>

Table 3.1. Parameters of the model of Alexander and Haasen.
deformation experiments for GaAs (Guruswamy et al., 1987) and InP (Müller et al., 1985) show different quantitatively behavior.

Tsai et al. (1992) reported estimates for the multiplication rate constant \( K \) and the strain hardening factor \( A \) for GaAs. Tsai et al. (1992) used the model of Alexander and Haasen to fit the experimental values of the upper and lower yield stresses from dynamic deformation of GaAs samples at \( T = 673 \) K and strain rate \( \dot{\varepsilon} = 1.8 \times 10^{-4} \) s\(^{-1} \) with initial dislocation density \( N_0 = 4 \times 10^4 \) cm\(^{-2} \) (Yonenaga et al., 1987). The values of \( K \) and \( A \) found by this procedure were \( 7 \times 10^{-3} \) m/N and 3.13 N/m, respectively (Tsai et al., 1992). According to these results \( K_{\text{GaAs}} \approx K_{\text{InP}} \). An estimation of the multiplication rate constant for GaAs will be given in Chapter 5, based on a constitutive theory for the critical resolved shear stress measured during the deformation of diamond/zincblende structure semiconductors. This estimate is based on data from deformation experiments for GaAs (Guruswamy et al., 1987) and InP (Müller et al., 1985) at \( T = 973 \) K and gives \( K_{\text{GaAs}} \approx 0.2K_{\text{InP}} \) with an accuracy of a factor of three. The conclusion from both these estimations is that the multiplication rate constant for GaAs is lower than the corresponding value for InP.

### 3.2.3. Plastic Deformation Dynamics of a Growing Crystal

The phenomenological model of Alexander and Haasen presented in Section 3.2.1 is generalized for three-dimensional deformation in a frame of reference that moves with the constant velocity of a growing crystal. Incorporation of dislocation convective fluxes as formulated for the derivation of the general macroscopic equation (3.12) for the evolution of the dislocation density in a growing crystal and ideal plasticity theory (Boley and Weiner, 1960) for the computation of the plastic strain field in the growing crystal are required to extend the equations appropriate for uniaxial deformation to the three-dimensional deformation in a growing crystal. The mathematical formulation for the calculation of the thermal elastoplastic stress field in the crystal is presented in Section 3.2.3.1 and the development of the dislocation dynamics model is presented in Section 3.2.3.2.
3.2.3.1. Mathematical Formulation for the Calculation of the Thermal Elastoplastic Stress Field

The temperature, stress, and strain fields developed in the crystal are assumed to vary with time only because of the rigid body translation of a volume element of the crystal because of crystal pulling in the growth direction. Under these quasi-steady state growth conditions there is only spatial variation in the crystal of temperature, stress and strain.

The stress \( \sigma(x) \) and strain \( \varepsilon(x) \) tensors evaluated at any point \( x \) are decomposed into isotropic \((-p \text{ and } \epsilon)\) and deviatoric \((S \text{ and } \varepsilon)\) parts as

\[
\sigma = -pI + S, \tag{3.19a}
\]

\[
\varepsilon = \varepsilon I + \varepsilon, \tag{3.19b}
\]

where \( I \) is the unit tensor, \(-p \equiv \frac{1}{3} \text{tr}(\sigma)\) and \( \epsilon \equiv \frac{1}{3} \text{tr}(\varepsilon)\) and the deviatoric parts are traceless. The isotropic stress and strain are assumed to behave according to classical linear thermoelasticity (Boley and Weiner, 1960) as

\[
\varepsilon(x) = \frac{p(x)(1-2\nu)}{E} + \alpha T(x), \tag{3.20}
\]

where \( \alpha \) is the coefficient of thermal expansion, \( \nu \) is Poisson's ratio, \( E \) is Young's modulus and \( T(x) \) is the temperature field in the crystal.

According to ideal plasticity theory (Boley and Weiner, 1960) the deviatoric part of the strain tensor \( \varepsilon \) is decomposed into elastic \((\varepsilon^E)\) and plastic \((\varepsilon^p)\) components as

\[
\varepsilon \equiv \varepsilon^E + \varepsilon^p. \tag{3.21}
\]

The strain is related to a fictitious deviatoric stress \( S^E \) that would be present if the material was perfectly elastic as

\[
\varepsilon \equiv \frac{1}{2G} S^E, \tag{3.22}
\]

where \( G \equiv \frac{E}{2(1+\nu)} \) is the shear modulus of the material. Likewise, the deviatoric stress \( S \) is assumed to result entirely from the elastic deformation \( \varepsilon^E \), so that

\[
S \equiv 2G\varepsilon^E. \tag{3.23}
\]
Combining equations (3.21)-(3.23) gives

$$\mathbf{S} = \mathbf{S}^E - 2\mathbf{G} \mathbf{e}^p,$$

(3.24)

where the first term on the right is the stress in the crystal if all the strain was ideally elastic and the second term is the correction for the plastic component in the real elastoplastic material. Equation (3.24) is the starting point for the analysis of plastic deformation in the three-dimensional crystal.

The stress field \( \varepsilon(\mathbf{x}) \) in the ideal elastic crystal with strain \( \varepsilon(\mathbf{x}) \) is computed by the solution of the linear thermoelastic problem formed by the constitutive equation

$$\varepsilon = \frac{1 + \nu}{E} \varepsilon - \frac{\nu}{E} (\text{tr}(\varepsilon)) \mathbf{I} + \alpha T \mathbf{I},$$

(3.25)

Cauchy’s law (Love, 1927; Boley and Weiner, 1960)

$$\nabla \cdot \varepsilon = 0,$$

(3.26)

and the compatibility conditions (Boley and Weiner, 1960)

$$\nabla \times \varepsilon \times \nabla = 0,$$

(3.27)

where the strain tensor is related to the displacement vector field by equation (2.1) or

$$\varepsilon \equiv \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^T],$$

(2.1)

where the notation \((\cdot)^T\) denotes the transpose of a second rank tensor.

**3.2.3.2. Model of Plastic Deformation Dynamics During Growth**

The equations of the model of Alexander and Haasen are formulated for a general volume element of crystal assuming that a sufficient number of dislocations exist in this element to define a dislocation density field \( N(\mathbf{x}) \). According to the quasi-steady state assumption introduced in Section 3.2.3.1 time variations occur only because of the rigid body translation of the volume element at speed \( V_p \) in the \( z \)-direction, as occurs because
of crystal translation with pull velocity $V_p\hat{z}$ in both the CZ and LEC methods. Then the material derivative (Bird et al., 1960) for the volume element is simply

$$\frac{D(\_)}{Dt} = V_p \frac{\partial (\_)}{\partial z}.$$ 

(3.28)

Introducing the material derivative in place of the time derivative in the Orowan equation (3.13) gives

$$\frac{De^p}{Dt} = \phi b v N(x),$$

(3.29)

where $e^p$ is a scalar invariant of the deviatoric plastic strain tensor $e^p$, and $\phi$ is a geometrical factor that relates shear to compression and tensile strain. The main differences between equation (3.29) and equation (3.13) which expresses the plastic deformation for a specific slip system is that a scalar invariant of the complete plastic strain tensor replaces the plastic strain component over a single slip system. The use of the material derivative of a scalar plastic strain instead of the material derivative of the complete plastic strain tensor for three-dimensional deformation is justified below by the relationship between the two material derivatives.

The geometrical factor $\phi$ depends on the growth direction and the orientation of the crystallographic planes (Hirth and Lothe, 1982) and expresses the projection operation which is performed when the stress tensor is resolved on a slip system for the calculation of the scalar resolved shear stress (Hirth and Lothe, 1982). In a similar manner, equation (3.18) is rewritten as

$$\frac{DN(x)}{Dt} = K_{\tau eff} v N(x).$$

(3.30)

Combining equations (3.29) and (3.30) with equations (3.14) and (3.17) gives a reduced one-dimensional form of the model of Alexander and Haasen for the plastic strain and the dislocation density as

$$\frac{De^p}{Dt} = \phi b B_0 \exp\left(-\frac{Q}{kT}\right)(\tau - A\sqrt{N})^m N,$$

(3.31)

$$\frac{DN}{Dt} = KB_0 \exp\left(-\frac{Q}{kT}\right)(\tau - A\sqrt{N})^{m+1} N,$$

(3.32)
for nonzero values of the scalar effective stress. It should be mentioned that equation (3.32) has exactly the form of the general macroscopic equation (3.12) for the evolution of the dislocation density. The convective flux of equation (3.12) has been incorporated into the material derivative and the multiplication and pair annihilation rates are given by

\[ M_N = K \tau v \ N, \quad (3.33) \]

\[ L_{2,N} = K A v \ N^{3/2}, \quad (3.34) \]

To complete the model requires defining the scalar stress \( \tau \) in terms of the stress tensor \( \sigma \) and generalizing eq. (3.31) to model the tensorial plastic strain \( \varepsilon^p \). Equation (3.24) gives the deviatoric stress tensor \( \mathbb{S} \) in terms of the deviatoric stress tensor for the perfectly elastic deformation and \( \varepsilon^p \). We define a scalar form of eq. (3.24) in terms of the second invariant of each tensor as

\[ S \equiv S^E - \frac{G}{\phi} \varepsilon^p, \quad (3.35) \]

where \( S \equiv \sqrt{\frac{1}{2} \mathbb{S} : \mathbb{S}}, \ S^E \equiv \sqrt{\frac{1}{2} \mathbb{S}^E : \mathbb{S}^E} \) and the scalar plastic strain \( \varepsilon^p \) is related to the tensor \( \varepsilon^p \) by

\[ \frac{D \varepsilon^p}{Dt} = \frac{1}{S} \left( \frac{D \varepsilon^p}{Dt} \right) \mathbb{S}. \quad (3.36) \]

Equation (3.36) expresses the von Mises yield condition (Boley and Weiner, 1960) according to which the time derivative of the deviatoric plastic strain tensor is proportional to the deviatoric stress tensor with a nonzero proportionality factor. Equating the second invariants of the two sides of eq. (3.36) and integrating in \( z \) shows that

\[ \varepsilon^p = \frac{1}{V_p} \int_0^z \sqrt{\frac{1}{2} \left( \frac{D \varepsilon^p}{Dt} : \frac{D \varepsilon^p}{Dt} \right)} \, dz + \varepsilon_0^p, \quad (3.37) \]

where \( \varepsilon_0^p \) is the plastic strain in the crystal at the melt/crystal interface. However, the scalar plastic strain which is used in the decomposition of equation (3.35) is defined as the square root of the second invariant of the deviatoric plastic strain tensor, i.e.

\[ \varepsilon^p \equiv \sqrt{\frac{1}{2} \mathbb{\varepsilon}^p : \mathbb{\varepsilon}^p}. \quad (3.38) \]
Derivation of the scalar equation (3.35) from the tensorial equation (3.24) is rigorous in the asymptotic limit of rapid dislocation multiplication that is analyzed in Section 3.3. The validity of the scalar form (3.35) with \( \varepsilon^p \) defined by equation (3.38) is demonstrated in Section 3.3.4.

The square root of the second invariant of the stress tensor is called the von Mises stress (Boley and Weiner, 1960). In this analysis, the term von Mises stress (\( \sigma \)) is used for the square root of the second invariant of the elastic stress tensor or \( \sigma \equiv S^E \).

The applied stress \( \tau \) is equated with \( S \) in eq. (3.35). Substituting (3.35) into eqs. (3.31) and (3.32) and using the definition of the material derivative (3.28) gives the multidimensional Haasen model as

\[
\frac{d\varepsilon_p}{dz} = \frac{\phi bB_0}{V_p} \exp\left(-\frac{Q}{kT}\right) (S^E)^m \left[1 - \frac{G\varepsilon_p}{\phi S^E} - \frac{A\sqrt{N}}{S^E}\right]^m N, \tag{3.39}
\]

\[
\frac{dN}{dz} = \frac{KB_0}{V_p} \exp\left(-\frac{Q}{kT}\right) (S^E)^{m+1} \left[1 - \frac{G\varepsilon_p}{\phi S^E} - \frac{A\sqrt{N}}{S^E}\right]^{m+1} N, \tag{3.40}
\]

\[
\frac{d\varepsilon_p}{dz} = \frac{1}{S} \left( \frac{d\varepsilon_p}{dz} \right) S, \tag{3.41}
\]

The equations (3.39)-(3.41), Cauchy’s law (3.26), and the compatibility conditions (3.27) for the components of the strain tensor completely define the stress and strain fields and the dislocation density in a crystal with given temperature distribution. Dillon et al. (1987) developed an iterative numerical method for solving this equation set. The application of this scheme to the calculation of axisymmetric fields for either CZ or LEC crystal growth would be as follows. The stress \( \sigma \) and elastic strain \( \varepsilon^E \) fields are computed from a first approximation to \( \varepsilon^n \) by solving Cauchy’s law with the compatibility conditions. The square root of the second invariant \( S^E \) is computed for each radial position \( r \) in the cylindrical crystal and the ordinary differential equations, (3.39) and (3.40), are integrated from values at the melt/crystal interface

\[
\varepsilon^p(r, 0) = \varepsilon^p_0, \quad N(r, 0) = N_0(r). \tag{3.42}
\]

The plastic strain tensor is constructed according to eq. (3.41) and the iteration is repeated using this new approximation to \( \varepsilon^n \) to compute stresses and strains until convergence is
achieved in the components of \( \varepsilon_p \). The numerical solution of either equation set is not pursued in this thesis. Instead, we examine asymptotic approximations for the plastic strain and dislocation density that result because of the operating conditions in CZ and LEC crystal growth. Here, the qualitative features of the elastic stress fields computed in previous analyses are used to infer the evolution of the dislocation density.

3.3. Scaling Analysis and Asymptotic Calculation of Dislocation Density

The asymptotic analyses hinge on analysis of the equations for the plastic strain \( \varepsilon_p \) and dislocation density \( N \) and are based on limits where the details of the temperature field and the resulting stress field are relatively unimportant. The key point of this analysis is that the qualitative features of the dislocation density for CZ and LEC crystal growth conditions, according to the model of Alexander and Haasen, is predicted by asymptotic analysis. The quantitative details of the elastic stress field throughout the crystal are not necessary to carry out the asymptotic analysis; however, these details must be known for the quantitative calculation of dislocation density profiles. As described in this section, the very rapid rate of formation of dislocations under typical Czochralski growth conditions leads to equilibration of the dislocation density with the local stress field after only a short development length, or boundary layer, near the melt/crystal interface. Because of this structure, the qualitative structure of the dislocation field can be estimated based on only the magnitude and gradient of the thermal elastic stress along the crystal length.

3.3.1. Scaling for Dislocation Density and Plastic Strain

The evolution of the dislocation density as a function of the magnitude of the plastic strain \( \varepsilon_p \) is governed by the ordinary differential equation formed by combining equations (3.39) and (3.40) as

\[
\frac{dN}{d\varepsilon_p} = \frac{K}{\phi b}(S^E - \frac{G_\varepsilon_p}{\phi} - A\sqrt{N}),
\]

(3.43)

where the first term in parenthesis is the elastic stress, the second is the stress relief caused by plastic strain and the third is the back stress due to the dislocation interactions.
Equation (3.43) suggests natural scales for each variable. We set \( \sigma_0 \) as the characteristic size of the von Mises stress \( \sigma \equiv S^E \) and define the scaled plastic strain \( x \), the dimensionless dislocation density \( y \) and the dimensionless von Mises stress \( \bar{\sigma} \) as

\[
x \equiv -\frac{Ge^p}{\phi \cdot \delta}, \quad y \equiv \frac{Gb}{K\sigma_0^2N}, \quad \bar{\sigma} \equiv \frac{\sigma}{\sigma_0}.
\]  

(3.44)

Equation (3.43) becomes in dimensionless form

\[
\frac{dy}{dx} = (\bar{\sigma} - x - \delta \sqrt{y}),
\]

(3.45)

where \( \delta \) is only a function of material parameters

\[
\delta \equiv \left[ \frac{A^2 K}{Gb} \right]^{1/2},
\]

(3.46)

and measures the size of the back stress relative to the elastic stress.

An important result of this scaling analysis is that \( \delta \) is small for the materials of interest here; \( \delta = 0.11 \) for InP and \( \delta = 0.019 \) for Si, using the parameter estimates given in Section 3.2.2. Neglecting the back stress for \( \delta \ll 1 \) greatly simplifies the analysis by decoupling the calculation of the dislocation density \( y \) and the plastic strain \( x \). This approximation is used below in calculations for specific axial stress profiles in the crystal.

Finally, we scale the axial distance \( z \) with the crystal radius \( R \) to give the dimensionless coordinate \( \zeta = z/R \). This scaling is appropriate for the variation of the temperature field and, hence, the elastic stress in the crystal.

### 3.3.2. Constant von Mises stress: \( \bar{\sigma}(\zeta) = 1 \)

When the von Mises stress \( \sigma \) and the temperature \( T \) are approximately constant in the region of dislocation generation we may take \( \bar{\sigma} = 1 \) and reduce eq. (3.45) to

\[
\frac{dy}{dx} = 1 - x + O(\delta),
\]

(3.47)
for $\delta \ll 1$. Assuming that the plastic strain is zero for a zero dislocation density, i.e., $y(0) = 0$, and integrating (3.47) gives

$$y = x - \frac{x^2}{2} + O(\delta),$$

or in dimensional form

$$N \simeq \frac{K \sigma_0}{\phi b} \epsilon_0 - \frac{K G}{2 \phi^2 b} (\epsilon_0)^2.$$ (3.49)

This quadratic expression is in qualitative agreement with the experimental results for germanium reported by Alexander and Haasen (1968); these results are shown in Fig. 3.4.

Introducing the dimensionless variables $(x, y, \zeta)$ into eqs. (3.39) and (3.40) with $\tilde{\sigma} = 1$ gives

$$\frac{dx}{d\zeta} = Da \left(1 - x - \delta \sqrt{y}\right)^m y,$$ (3.50)

$$\frac{dy}{d\zeta} = Da \left(1 - x - \delta \sqrt{y}\right)^{m+1} y,$$ (3.51)

where the dimensionless group

$$Da \equiv \frac{KB_0 R}{V_p} \exp\left(-\frac{Q}{kT}\right) \sigma_0^{m+1}$$ (3.52)

is called the Dislocation Damköhler Number, in a form analogous to the Damköhler Number for a tubular chemical reactor (Buckmaster, 1985). The constant $Da$ scales the time scale for generating dislocations by the time scale for motion of the crystal through the region with applied stress $\sigma_0$ and temperature $T$. For $Da \ll 1$, dislocations are formed slowly relative to the motion of the crystal. For $Da \gg 1$, dislocations are formed rapidly and the dislocation density causes rapid plastic stress relief, which eventually halts the generation dislocations. A thin boundary layer in $\zeta$ is expected in this limit in which $x(\zeta)$ and $y(\zeta)$ vary from the initial values at $\zeta = 0$

$$y(0) = y_0, \quad x(0) = x_0$$ (3.53)

to the equilibrium values for $\zeta = O(1)$.
Figure 3.4. Dependence of the local dislocation density on local strain after compression creep of germanium samples at $T = 855$ K under stresses of (a) 0.25 kg/mm$^2$ and (b) 0.5 kg/mm$^2$ (Alexander and Haasen, 1968).
The Dislocation Damköhler Number is very large for conditions typical for CZ and LEC crystal growth. For InP, picking \( T = T_m, R = 1.5 \text{ cm} \) and \( V_p = 1 \text{ cm/h} \) gives

\[
Da \approx \begin{cases} 
150, & \text{if } \sigma_0 = 0.1 \text{ MPa}; \\
37050, & \text{if } \sigma_0 = 1 \text{ MPa}, 
\end{cases}
\]

where \( \sigma_0 \) is the von Mises stress next to the melt/crystal interface; the value 1.0 MPa is approximately equal to the CRSS of InP measured at \( T \approx 0.8T_m \) (Müller et al., 1985). For Si, picking \( T = T_m, R = 10 \text{ cm} \) and \( V_p = 10 \text{ cm/h} \) gives

\[
Da \approx \begin{cases} 
1000, & \text{if } \sigma_0 = 1 \text{ MPa}; \\
29000, & \text{if } \sigma_0 = 5 \text{ MPa}. 
\end{cases}
\]

The boundary layer in the dislocation density and plastic strain is best seen by setting \( \varepsilon \equiv Da^{-1} \) and rewriting eqs. (3.50) and (3.51) as

\[
\varepsilon \frac{dx}{d\zeta} = (1 - x - \delta \sqrt{y})^m y, \quad (3.54)
\]

\[
\varepsilon \frac{dy}{d\zeta} = (1 - x - \delta \sqrt{y})^{m+1} y. \quad (3.55)
\]

For \( \varepsilon \ll 1 \), eqs. (3.54) and (3.55) have solutions with a boundary layer adjacent to the melt/crystal interface (\( \zeta = 0 \)) of dimensionless thickness \( O(\varepsilon) \) in which the temperature is essentially \( T_m \) and the assumption of constant stress is valid.

Outside the boundary layer, the dislocation and plastic strain fields are given by the quasi-equilibrium conditions obtained by setting \( \varepsilon \to 0 \) in eqs. (3.54) and (3.55). These equations reduce to the single condition

\[
1 - x - \delta \sqrt{y} = 0. \quad (3.56)
\]

Using the assumption of negligible back stress (\( \delta \ll 1 \)) yields the quadratic expression (3.48) for the dislocation density, which is inverted to solve for \( x \) as

\[
x = 1 \pm \sqrt{1 - 2y} + O(\delta). \quad (3.57)
\]
Because $e^n$ is an increasing function of the dislocation density, only one solution of (3.57) makes physical sense, i.e.

$$x = 1 - \sqrt{1 - 2y} + O(\delta).$$  \hspace{1cm} (3.58)

The maximum value of the dislocation density is given by equation (3.56) as

$$y = \frac{1}{2} + O(\delta),$$  \hspace{1cm} (3.59)

or in dimensional form as

$$N_\infty \equiv N(\infty) \simeq \frac{K\sigma_0^2}{2Gb}.$$  \hspace{1cm} (3.60)

The value of the dislocation density $N_\infty$ occurs for the given stress level $\sigma_0$ after a small development length of size $O(\varepsilon)$.

The dislocation density inside the boundary layer is computed by substituting eq. (3.58) into (3.55) for $\delta \ll 1$ to give

$$\varepsilon \frac{dy}{d\zeta} = (1 - 2y) \frac{m+1}{2} y + O(\delta).$$  \hspace{1cm} (3.61)

Equation (3.61) has exact solutions only when $(\frac{m+1}{2})$ is an integer. This exponent is approximately equal to 1.05, 1.2 and 1.35 for silicon, InP and GaAs, respectively. Hence, the solutions of eq. (3.61) for $\frac{m+1}{2}$ equal to one and two are upper and lower bounds for these materials. These solutions are:

$$y(\zeta) = \frac{y_0 e^{D\zeta}}{1 + 2y_0(e^{D\zeta} - 1)},$$  \hspace{1cm} (3.62a)

for $(m + 1)/2 = 1$, and

$$D\zeta = \ln \left[ \frac{y(1 - 2y_0)}{y_0(1 - 2y)} \right] + \frac{2(y - y_0)}{(1 - 2y_0)(1 - 2y)},$$  \hspace{1cm} (3.62b)

for $(m + 1)/2 = 2$. The profiles for InP with $\sigma_0 = 0.1$ MPa and $V_p = 1$ cm/hr are shown in Fig. 3.5 in dimensional form with $y_0 = 7.8 \cdot 10^{-4}$ ($N_0 = 1$ cm$^{-2}$). The boundary layer

105
Figure 3.5. Variation of the dislocation density with the distance from the melt/crystal interface predicted by equations (3.62a) and (3.62b). Results are for InP at a constant von Mises stress of 0.1 MPa. The solid (-----) and dashed (· · ·) curves correspond to $m = 1$ and $m = 3$, respectively. The above curves provide upper and lower bounds for the axial variation of the dislocation density for Si ($m = 1.1$), GaAs ($m = 1.7$) and InP ($m = 1.4$).
spans between \( N_0 \) and \( N_\infty \approx 605 \) dislocations/cm\(^2\). The variation caused by selecting \( 1 < m < 2 \) is significant only in the transition at the outer edge of the boundary layer.

Finally, the effects of the back stress on the predictions of the final dislocation density and plastic strain are estimated by expanding \( x \) and \( y \) in a power series in \( \delta \). The leading order corrections are:

\[
x = 1 - \frac{\delta}{\sqrt{2}} + O(\delta^2)
\]  
(3.63a)

for the strain and

\[
y = \frac{1}{2} + O(\delta)
\]  
(3.63b)

for the dislocation density.

3.3.3 Spatially Varying von Mises Stress: \( \bar{\sigma} = \bar{\sigma}(\zeta) \)

During crystal growth, both the temperature and the elastic stress in the crystal typically vary with axial position. We account for these variations by allowing \( \bar{\sigma} = \bar{\sigma}(\zeta) \) and a scaled temperature \( \bar{T}(r, \zeta) \equiv T(r, \zeta)/T_m \) to vary with \( \zeta \). Then the Boltzmann factor in the Dislocation Damköhler Number (3.52) is factorized as

\[
\exp\left(-\frac{Q}{kT}\right) = af(\bar{T}),
\]  
(3.64)

where

\[
f(\bar{T}) \equiv a^{(1-T(\zeta))/T(\zeta)}, \quad a \equiv \exp\left(-\frac{Q}{kT_m}\right).
\]  
(3.65)

Then \( Da \) is defined as

\[
Da \equiv \varepsilon^{-1} = \frac{KB_0 R a \sigma_0^m + 1}{V_p}
\]  
(3.66)

and the dimensionless equations for the Haasen model reduce to

\[
\varepsilon \frac{dx}{d\zeta} = f(\bar{T})(\bar{\sigma} - x - \delta \sqrt{y})^m y,
\]  
(3.67)

\[
\varepsilon \frac{dy}{d\zeta} = f(\bar{T})(\bar{\sigma} - x - \delta \sqrt{y})^{m+1} y.
\]  
(3.68)
Two cases are treated separately. In the first, the von Mises stress $\bar{\sigma}$ is taken as a *decreasing function of axial distance*. This is typically the case in CZ growth of silicon, where the largest radial temperature differences exist at the melt/crystal interface (Brown, 1989). The stresses decrease with axial distance over a length scale characteristic of the temperature field. The crystal radius characterizes this variation and the gradient $d\bar{\sigma}/d\zeta$ is $O(1)$.

In the second case, the largest elastic stress in the crystal exists at some distance away from the interface and *the stress increases with axial distance from the interface*. This is the situation in LEC growth, where the largest stresses exist at the top of the encapsulant (Duceax, 1983; Völkl and Müller, 1987). Because the encapsulant thickness is the same order-of-magnitude as the crystal radius, $d\bar{\sigma}/d\zeta = O(1)$.

In both cases, for axial temperature gradients that are typically between 10 and 50 K/cm, the temperature function $f(\bar{T})$ varies between $1/3 < \bar{T} < 1$ and so does not change the scaling of the dislocation boundary layer.

3.3.3.1 Decreasing von Mises Stress: $d\bar{\sigma}/d\zeta < 0$

If the length scale for the variation of the dimensionless von Mises stress $\bar{\sigma}(\zeta)$ is big relative to the length scale for the dislocation boundary layer $\varepsilon \equiv Da^{-1}$, the equilibrium value of the dislocation density $N_\infty$ is given by equation (3.60), with $\sigma_0$ taken as the interfacial value of the von Mises stress.

This solution structure can be seen rigorously by solving eqs. (3.67) and (3.68) with $f(\bar{T}) = \text{constant}$, and $\bar{\sigma} = 1 - g_\sigma \zeta$, where $g_\sigma$ is the dimensionless gradient of the stress. An exact solution is obtained in terms of exponentials and error functions in the limit $\delta \ll 1$ and $m = 1$. This solution has exactly the same boundary layer structure described in Section 3.3.2 and leads to the same final value of the dislocation density $N_\infty(r)$, eq. (3.60).

The decreasing stress with axial distance leads to internal stresses that are higher than the applied stress and, therefore, to zero effective stress for dislocation generation. The dislocation density remains fixed at $N_\infty(r)$. 

108
3.3.3.2 Increasing von Mises Stress: $d\bar{\sigma}/d\zeta > 0$

For $Da \gg 1$, the boundary layer in the dislocation density occurs on a length scale $O(Da^{-1})$ small relative to the variation in the von Mises stress, which is assumed to vary as $\bar{\sigma}(\zeta) = 1 + g_\sigma \zeta$, where $g_\sigma$ is positive and $O(1)$. If $T \simeq T_m$ and $f(\bar{T}) \simeq 1$ inside this layer, eqs. (3.67)-(3.68) reduce to eqs. (3.50)-(3.51) and give the same boundary layer structure described in Section 3.3.2.

The increasing von Mises stress does affect the solution away from the boundary layer. In this region $\epsilon \ll 1$ in eqs. (3.67)-(3.68) leads to

$$\bar{\sigma}(\zeta) - x(\zeta) - \delta \sqrt{y(\zeta)} = 0. \quad (3.69)$$

For negligible back stress ($\delta \ll 1$), eq. (3.69) gives

$$x = \bar{\sigma}(\zeta) + O(\delta), \quad (3.70)$$

which gives the scaling between the plastic strain and the von Mises stress far from the interface.

Eqs. (3.67) and (3.68) are combined into the single differential equation (3.45) for $y(x)$. An explicit relation between $x$ and $\bar{\sigma}$ is necessary to integrate eq. (3.45) for $y(x)$ and eq. (3.70) does not provide such a relation. Equation (3.69) gives sufficient information for the derivation of an upper bound for the dislocation density $y$. The result for constant stress (3.60) guarantees that $\sqrt{y} < \bar{\sigma}$, even in the case where the entire crystal feels the maximum stress level. Therefore, eq. (3.69) yields the condition $1 - \frac{x}{\bar{\sigma}} < \delta$, or equivalently

$$\bar{\sigma} < \frac{x}{1 - \delta}. \quad (3.71)$$

Taking the upper bound on the stress as

$$\bar{\sigma} \simeq \frac{x}{1 - \delta}, \quad (3.72)$$

and substituting into eq. (3.45) gives

$$\frac{dy}{dx} \simeq \frac{x}{1 - \delta} - x - \delta \sqrt{y}. \quad (3.73)$$

109
Expanding eq. (3.73) in powers of $\delta$ gives to leading order the equation

$$\frac{dy}{dx} \simeq \delta x,$$  \hspace{1cm} (3.74)

which is integrated to give

$$y \simeq y_i + \frac{\delta}{2}(x^2 - x_i^2),$$  \hspace{1cm} (3.75)

where $y_i$ and $x_i$ are the initial values of the dislocation density and the plastic strain in the bulk region away from the interface. Using the results of eqs. (3.63a) and (3.63b) for these initial values and expressing $x$ from eq. (3.72) gives

$$y \simeq \frac{1}{2} + \frac{\delta}{2} \left[ \tilde{\sigma}^2 (1 - 2\delta) - 1 \right] + O(\delta^2).$$  \hspace{1cm} (3.76)

Equation (3.76) provides the scaling $y \simeq \delta \tilde{\sigma}^2$ of the dislocation density in the case where the von Mises stress increases with the distance above the melt/crystal interface. In dimensional form, this scaling is

$$N_s(R\zeta) = \delta \frac{K\sigma^2(R\zeta)}{Gb}.$$  \hspace{1cm} (3.77)

and dictates the influence of the variation of the von Mises stress on the dislocation density in the bulk crystal. If $\sigma$ increases with $\zeta$, the dislocation density increases proportional to the square of this variation.

The scalings for the variables, $x = O(\tilde{\sigma})$ and $y = O(\delta \tilde{\sigma}^2)$, in the bulk are the starting point for a matched asymptotic expansion solution for the combined boundary layer and outer region problems, where the variables are expanded in power series of $\varepsilon$. No benefit is derived from this calculation, because the rescaled equations governing $x$ and $y$ in the boundary layer must be solved numerically, just as the original equations, (3.67) and (3.68). The most interesting results hinge on the importance of the scalings for CZ and LEC crystal growth; these are discussed in Section 3.4.
3.3.4 Accuracy of Asymptotic Results

We have used the parameters of InP and conditions typical of LEC growth as a test of the accuracy of our asymptotic predictions against the numerical solutions of eqs. (3.67) and (3.68). An Explicit Euler integration method (Dahlquist and Björck, 1974) with appropriate control of the axial step size is sufficient for these calculations. Because of the steep boundary layer in the dislocation density expected near the melt/crystal interface the first step size was less than 1 μm in size. The value of the initial dislocation density \( N_0(r,0) \) only affects the dislocation density in the boundary layer and any value between zero and the dislocation density in the bulk suffices.

The initial value of the plastic strain was computed from the integration of eq. (3.43) with constant applied stress \( \tau \equiv S \). Because the plastic contributions are negligible at \( z = 0 \), the von Mises stress \( \sigma_0 \) is taken to be equal to the applied stress. This integration, with the initial condition \( N = 0 \) at \( \varepsilon^p = 0 \) gives the expression

\[
\varepsilon^{p,0} = 2\phi \frac{b\sigma_0}{K A^2} \left[ - \frac{A\sqrt{N_0}}{\sigma_0} - \ln \left( 1 - \frac{A\sqrt{N_0}}{\sigma_0} \right) \right].
\]

(3.78)

for the initial plastic strain.

For constant or decreasing von Mises stress \( \bar{\sigma} \), the numerical results are qualitatively the same as the asymptotic results depicted in Figure 3.5 for \( \sigma_0 = 0.1 \) MPa. A direct comparison between the numerical computations and the asymptotic results is shown in Figure 3.6 for \( \sigma_0 = 0.1 \) Mpa. The results are in excellent quantitative agreement and differ numerically because of the correction caused by neglecting the back stress \( (\delta \ll 1) \) in the asymptotic analysis.

The asymptotic results for a von Mises stress that increases with axial distance are compared to numerical solution of eqs. (3.67) and (3.68) for \( \sigma(R \zeta) = \sigma_0 + G_\sigma R \zeta \) over a length \( \Delta = L/R \) that is meant to model the height of the encapsulant in an LEC crystal growth system. We model the variation in the von Mises stress by setting the gradient \( G_\sigma \) so that the stress varies from 0.1 MPa at the interface to 1.0 MPa at the encapsulant layer at \( \Delta = 2.0 \) and a crystal with radius 1.5 cm. For such a variation in the von Mises stress, the details of the temperature profile affect only the temperature function \( f(\bar{T}) \).

111
Figure 3.6. Comparison between the asymptotic results of eqs. (3.62a) and (3.62b) and the numerical solution of eqs. (3.67) and (3.68) for InP in the cases of a constant or decreasing von Mises stress with the distance above the melt/crystal interface: \( \sigma_0 = 0.1 \) MPa.
The comparison between the prediction of the asymptotic expression eq. (3.70) and the numerical calculation for the plastic strain $\varepsilon^p$ shown in Fig. 3.7 clearly demonstrates the validity of the scaling arguments and the usefulness of the asymptotic results. The numerical calculations are performed for a constant temperature gradient of -50 K/cm. Outside the thin boundary layer that is captured by the numerical calculations, the profiles of $\varepsilon^p$ are essentially identical. These profiles are independent of the initial dislocation density and the initial plastic strain, eq. (3.78).

The asymptotic expression eq. (3.76) is compared with numerical results for the dislocation density for two different temperature profiles in Fig. 3.8 to understand the effect of the temperature function $f(\bar{T})$ on the dislocation density profile. The first calculation is performed for a constant temperature gradient -50 K/cm and the second for a temperature profile given by the fin approximation (Bird et al., 1960; Kobayashi and Iwaki, 1985) using typical values for the Peclet and Biot numbers according to the data given by Völkl and Müller (1987) for an estimate of the heat transfer coefficient. The temperature gradient predicted by the fin approximation is lower than the constant gradient everywhere in the crystal.

It is obvious that the temperature function affects the dislocation density profile. The lower temperature gradient for the fin approximation leads to higher predictions of the dislocation density from the Haasen model. The dislocation densities are the same order-of-magnitude with both temperature profiles and lower than the upper bound provided by the asymptotic result, eq. (3.70). The importance of the asymptotic result is that it gives a closed-form expression for the variation of the dislocation density far from the boundary layer of the same order-of-magnitude as the values computed by solution of the entire model.

Another point that should be discussed is the asymptotic validity of the scalar applied stress formulation. Equation (3.35) that has been used in the analysis, where the scalar plastic strain $\varepsilon^p$ is given by eq. (3.38) is exact to within $O(\delta)$ in the asymptotic limit $Da \gg 1$ and $\delta \ll 1$. Rapid dislocation multiplication expressed by $Da \gg 1$ guarantees complete plastic stress relief in the bulk crystal, i.e. outside of the boundary layer. The dislocation density increases inside the boundary layer from the initial value $N_0$ at $z = 0$
Figure 3.7. Variation of the plastic strain with the distance from the melt/crystal interface. The solid line is the numerical solution of eqs. (3.67) and (3.68) using a constant temperature gradient of -50 K/cm and a linear axial variation of the von Mises stress from 0.1 MPa at the interface to 1 MPa at a distance from the interface equal to the encapsulant height. The curve (---) is the asymptotic result of eq. (3.70).
Figure 3.8. Axial variation of the dislocation density for a linear axial variation of the von Mises stress with minimum value $\sigma_0 = 0.1$ MPa and maximum value $\sigma_{max} = 1$ MPa. The solid curve (-----) is the asymptotic result eq. (3.76) and the two curves are the results of numerical solution of eqs. (3.67) and (3.68) for a constant temperature gradient (- - - -) and the fin approximation (---) to the temperature profile.
to the bulk value expressed by the single condition (3.69) in the outer region. This condition is rewritten as

$$\frac{S}{\sigma_0} = \delta \sqrt{y}. \quad (3.79)$$

Therefore, $S/\sigma_0 \simeq 0$ in the outer region with $O(\delta)$ accuracy, which means that all the elements $S_{ij}$ of the deviatoric stress tensor are very small compared to $\sigma_0$ in the outer region. Combining eq. (3.79) with eq. (3.24) in component form gives

$$\frac{S_{ij}^E}{\sigma_0} = \frac{2G}{\sigma_0} e_{ij}^p + O(\delta), \quad (3.80)$$

which means that for each component of the stress and strain tensor the ratio $S_{ij}^E/2Ge_{ij}^p$ is a constant to within $O(\delta)$. The necessary and sufficient condition for the second invariant of the tensor $S$ to be equal to the square of the right-hand side of eq. (3.35) is that $S_{ij}^E/2Ge_{ij}^p$ is a constant. This is easily shown by taking the second invariant of the right-hand side of equation (3.24). This condition validates the use of equation (3.35) to the same degree of accuracy, $O(\delta)$ as the results of the asymptotic analysis.

It was noted in Section 3.2.3.2 that numerical procedures for simultaneous calculation of the dislocation field $N(x)$ and the stress $S$ and plastic strain $e^p$ fields requires initial approximations for the elements $e_{ij}^p$ of the deviatoric plastic strain tensor and iteration until convergence is achieved. The asymptotic validity of the scalar equation (3.35) guarantees the simultaneous calculation of the scalar stress $S$ and plastic strain $e^p$ by solving the ordinary differential equations (3.39) and (3.40) without using initial approximations for the components of the plastic strain tensor. Therefore, no iteration is required. Using the tensor equation (3.24) and solving eq. (3.41) gives the elements $e_{ij}^p$ of the deviatoric plastic strain tensor. Finally, equation (3.24) gives the elements $S_{ij}$ of the deviatoric stress tensor. The accuracy in the calculation of the elements $S_{ij}$ is determined by the accuracy in the calculation of the elements $e_{ij}^p$ of the plastic strain tensor, because the elements of the elastic components $S_{ij}^E$ are known exactly from the solution of the thermal elastic problem prior to the analysis of plastic deformation. The $O(\delta)$ accuracy in the calculation of $de^p/dz$ and $S$ from the asymptotic analysis in the bulk crystal guarantees the same
degree of accuracy in the calculation of the components $e_{ij}^p$ from eq. (3.41). Therefore, the entire problem for the thermal elastoplastic stress and strain fields is solved in the bulk crystal at the same degree of accuracy, $O(\delta)$, as the results of the asymptotic analysis.

3.4. Implications for Czochralski and Liquid Encapsulated Czochralski

Crystal Growth

The results derived in Section 3.3 have important implications for the growth of semiconductor materials by the CZ and LEC methods. These expressions are particularly useful because conclusions about the shapes of the dislocation profiles in crystals can be reached solely from examining the corresponding profiles of the von Mises stress. These conclusions are discussed here for the growth of InP and GaAs by the LEC technique and Si by CZ crystal growth.

3.4.1 Growth of InP and GaAs

Völkl and Müller (1989) pioneered the use of the model of Alexander and Haasen (1968) for analysis of dislocation production in InP growth by the LEC method. Their measurements of the parameters for the model make it possible to compare our analysis directly to their numerical calculations of the dislocation density profile based on simulations of the temperature and stress profiles in the crystal. The asymptotic analysis gives important interpretations to their numerical calculations.

As described in Section 3.3.4, these calculations show that in LEC growth with deep encapsulant layers the maxima in the von Mises stress in the crystal occur at the level of the crystal adjacent to the top of the encapsulant layer, not at the melt/crystal interface. Similar results have also been reported by Duseaux (1983) and Motakef and Witt (1987). In many of the simulations performed by Völkl and Müller (1989), the von Mises stress decreased from the value at the melt/crystal interface ($\zeta = 0$) before beginning to climb to a maximum at the top of the encapsulant ($\zeta = \Delta$). The initial decrease in the stress profile locks the dislocation density at the value given by eq. (3.60) that corresponds to the bulk value for the von Mises stress at the melt/crystal interface $\sigma = \sigma_0$. As the von Mises
stress increases above this interfacial value farther up the crystal, the dislocation density increases according to the relationship, eq. (3.76) for the bulk dislocation density in an increasing stress field. The dislocation density is maximum at the top of the encapsulant where the stress is highest.

It is well known empirically, that the dislocation density in LEC grown crystals of either InP or GaAs can be controlled by decreasing the maximum stress in the crystal either by changing the thermal environment or by altering the depth of the encapsulant to lower the magnitude of the thermal stresses in the crystal (Jordan et al., 1981; Motakef and Witt, 1987; Thomas, 1988). This effect is modeled here by examining the effect on the maximum dislocation density $N_{\text{max}}$ of lowering the maximum stress seen either at the top of the encapsulant or next to the melt/crystal interface. Results are presented for two LEC systems. In the first the encapsulant thickness is small so that high radial temperature gradients exist at the interface (Thomas, 1988). In this case, the maximum von Mises stress occurs next to the melt/crystal interface (Völkl and Müller, 1987; Thomas, 1988) and the dislocation density in the bulk crystal is given by the asymptotic result, eq. (3.60). In the second system the encapsulant is deep and the calculation is carried out following the analysis of Section 3.3.4 for increasing von Mises stress and for $\sigma_0 = 0.1$ MPa. In this case, the maximum von Mises stress occurs at the level of the crystal adjacent to the top of the encapsulant layer and the dislocation density in the bulk crystal is given by the asymptotic result, eq. (3.76). The results of the asymptotic analysis for both cases are presented in Fig. 3.9 demonstrating that there is a qualitative difference in the scalings of the dislocation density with stress that depends on the form of the stress profile. The result $N_{\text{max}} \propto \delta \sigma_{\text{max}}^2$, eq. (3.77) where $\delta \ll 1$ leads to lower values of bulk dislocation densities in the system with the deep encapsulant.

The results of the asymptotic analysis for the two LEC growth systems are directly compared with the numerical simulations of Völkl and Müller (1989), who also examined the effect on the maximum dislocation density of the reduction of the maximum von Mises stress in the crystal. The comparison is presented in Fig. 3.9. The asymptotic results are in quantitative agreement with the numerical results. In particular, the asymptotic results are always within 50% of the predictions of the numerical simulations in spite of the
Figure 3.9. Comparison between asymptotic (curves) and numerical results (○, □) for analysis of dislocation multiplication in LEC grown InP with shallow and deep encapsulant layer respectively, based on the model of Alexander and Haasen.
idealized stress profiles used in the asymptotic analysis as described in Section 3.3.4. In addition, the numerical simulations of Völk and Müller (1989) predict the experimentally measured etch pit densities in InP crystals grown by both LEC systems to within 50\% (Völk and Müller, 1989). This is a fairly good quantitative agreement considering the statistical variations of nearly half an order-of-magnitude of the etch pit densities within different growth runs under the same conditions. These comparisons show that the model of Alexander and Haasen and the asymptotic analysis lead to a direct interpretation of the influence of the stress gradient in the crystal on the dislocation density.

The asymptotic analysis also predicts the radial variations in the dislocation density observed in InP (Völk and Müller, 1989) and GaAs (Thomas, 1988) growth. In these systems the von Mises stress exhibits either a W-shaped or a U-shaped radial profile, as it is clearly demonstrated by Völk and Müller (1987) and Duseaux (1983) for the LEC growth of InP and GaAs, respectively. A schematic diagram of radial profiles of the von Mises stress and the dislocation density for InP or GaAs crystals grown by the LEC method is presented in Fig. 3.10. The axial and radial profiles of the von Mises stress strongly depend on the thermal boundary conditions. The LEC growth of InP with a deep encapsulant layer is an excellent example. In this system, the radial profile of $\sigma$ exhibits a minimum at $r = 0.6R$, a maximum at the periphery of the crystal and an intermediate value at the axis of the crystal (Völk and Müller, 1987). According to the asymptotic result, eq. (3.76), $N(r, z)$ is proportional to $\sigma^2(r, z)$ and the dislocation density is minimum at $r = 0.6R$, maximum at the periphery and intermediate at the center for some specific distance $z$ from the melt/crystal interface. This is in agreement with etch-pit measurements which show dislocation densities that are highest at the periphery of the crystal, intermediate at the center and lowest along an internal annulus (Jordan et al., 1980); the qualitative features of the etch pit density distribution were shown in Fig. 2.12. This qualitative agreement between etch-pit pattern and the stress profiles was an important foundation of the conclusions of Jordan et al. (1980) for relating dislocation generation to the CRSS in the material.

The dependence of the dislocation density on the direction of growth has been investigated by several authors (Jordan et al., 1980; Kobayashi and Iwaki, 1985; Lambropoulos,
Figure 3.10. Schematic diagram of the radial profile of (a) the von Mises stress and (b) the dislocation density in crystals of InP or GaAs grown by the LEC method.
1987). The orientation dependence enters the Haasen model by the geometrical factor $\phi$ in the Orowan equation (3.29). In our calculations we have used $\phi = 0.467$ as given by Völkl and Müller for (111) growth of InP. It is clear from the scales of eq. (3.44) and the asymptotic results (3.63a), (3.63b), (3.70) and (3.76) that the leading order term in the expression for the final plastic strain is proportional to $\phi$, while the leading order term in the expressions for the final dislocation density, eqs. (3.60) and (3.76) does not depend on $\phi$. This result does not contradict experimental data. Crystal orientation expressed by $\phi$ has an $O(\delta)$ effect on the final dislocation density which does not appear in the leading order terms of the asymptotic results (3.60) and (3.76). However, the dependence of dislocation density on $\phi$ is explicitly given by eq. (3.49). Resolving the stress tensor on the different slip systems and using the model of Alexander and Haasen (1968) for each slip system would provide the details of the effects of orientation on the dislocation density distribution in the crystal (Völkl, 1990).

There are important distinctions between the predictions of the Haasen model and the linear elastic theory based on the CRSS threshold. The correlation to the CRSS assumes that the favorite path for stress relief in the crystal is not plastic yield and that most of the stress is modelled by linear thermal elasticity. The Haasen model is based on *almost complete plastic stress relief* caused by the generation, propagation and multiplication of dislocations and lattice hardening caused by dislocation interactions. The quantitative and qualitative agreement between the Haasen model and measurements for the InP experiments of Völkl and Müller (1989) support the plastic mechanism of stress relief. Also, the fact that the thermal stresses predicted in GaAs (Jordan et al., 1980; Thomas, 1988; Motakef and Witt, 1987) and InP (Völkl and Müller, 1987, 1988) crystal growth systems exceed the CRSS by almost an order-of-magnitude imply that neglecting plastic stress relief is highly doubtful.

### 3.4.2 CZ Growth of Si

The dislocation density in Si crystals grown by the CZ method can be estimated directly using the parameters for the Haasen model given in Table 3.1 and typical operating
conditions for CZ crystal growth. Calculations of thermoelastic stresses in silicon crystal growth (Thomas, 1988; Kinney, 1992) show that, under typical operating conditions, the highest stresses appear at the melt/crystal interface. As a model calculation we take this stress to be equal to the CRSS of silicon at its melting temperature, \( \sigma_0 = 1.8 \text{ MPa} \). For this applied stress and at the melting temperature, the assumptions of negligible back stress (\( \delta \ll 1 \)) and high Dislocation Damkohler Number (\( Da \gg 1 \)) are valid; the computed values of these parameters are \( Da = 3390 \) and \( \delta = 0.019 \).

The asymptotic analysis for a crystal with its highest von Mises stress near the interface predicts that the dislocation density should reach its maximum value at the edge of the boundary layer there and should be proportional to \( \sigma_0^2 \), according to eq. (3.60). This prediction leads to a bulk dislocation density of \( N = 3760 \text{ cm}^{-2} \). However, CZ grown silicon is dislocation free! Clearly, either the mechanism of plastic stress relief modelled by the Haasen model must not be appropriate for silicon or so few dislocations form at the melt/crystal interface that the continuum description provided by the model is inappropriate.

Several causes for this discrepancy are possible. The model of Alexander and Haasen could be wrong. However, it describes well the influence of stress on dislocation generation in samples with initial dislocation densities (Alexander and Haasen, 1968). Possibly the temperature field in the crystal during CZ growth of Si may be so mild that the elastic stresses remain so far below the CRSS that dislocations can not multiply. This is the rational used by Jordan et al. (1980), who noted the higher values of the thermal conductivity and CRSS of silicon relative to III-V materials. These authors assumed that these differences are enough to lower the elastic stresses during silicon growth to the point that the material remains dislocation free. Examination of the multiplication rate constant leads to the same type of reasoning on the basis of the model of Alexander and Haasen, because \( K \) is two orders of magnitude smaller in silicon than in InP.

While both explanations may be feasible for growth of small diameter crystals \( (R < 2") \), neither can explain the growth of large-diameter dislocation free silicon. Very detailed modelling of large-diameter silicon growth indicates that the thermal stresses experienced by the crystal in a typical CZ system are in excess of the accepted value of
the CRSS threshold. This observation leaves only one possible explanation. The lack of dislocations in large-diameter silicon results because of inhibition to the nucleation of dislocations near the melt/crystal interface. This picture is consistent with the model of Alexander and Haasen and with the experience for growth of silicon (Lin and Benson, 1987).

If the initial dislocation density used in the Haasen model is \( N_0 = 0 \), the grown crystal will be dislocation free, irrespective of the level of the elastic stress. This prediction agrees with the use of the Dash method of seeding for silicon (Dash, 1959), which is necessary to produce dislocation free material. However, the model of Alexander and Haasen is not useful in this limit, because the dislocation multiplication law, eq. (3.18), relies on the existence of a dislocation density for multiplication and does not take into account the generation of new mobile dislocation loops from microdefects in the crystal (De Kock et al., 1975). This failure is obvious from the logarithmic singularity that is predicted by eq. (3.18) in the limit \( N \to 0 \). The important implication is that the effective prediction of the loss of dislocation free growth for large-diameter silicon crystals must involve an understanding of the mechanisms and rates for dislocation formation from various types of microdefects.

### 3.4.3 Effects of Necking During Seeding: Si versus InP

One possibility for the successful growth of dislocation free silicon is that the Dash method (Dash, 1959) for seeding effectively removes all dislocations that are created during the initial stages of growth and no other mobile dislocations form at the interface during bulk crystal growth. This mechanism focuses attention on the interesting point of the analysis that the results for the bulk dislocation density at high Dislocation Damköhler Number \( Da \gg 1 \) are independent of the crystal growth rate which only appears in \( Da \). This results implies that variations in the growth rate only affect the dislocation density through changes in the temperature and thermal stress fields in the crystal, not by changing the dynamics of dislocation formation, as long as conditions are still appropriate for assuming \( Da \) is large. Normally, increasing the crystal growth rate leads to steeper temperature
gradients in the crystal (Derby and Brown, 1986; Brown, 1988) and to high thermal stresses, thereby increasing the dislocation density.

The thin crystalline neck created in the Dash method by growing silicon at a growth rate in excess of 60 cm/h creates conditions where the assumption $Da \gg 1$ is no longer valid because of both the larger growth rate and the greatly decreased von Mises stresses associated with the thin crystal. To compare the effect of necking in silicon and InP growth, the Dislocation Damköhler Number is calculated in each case. For silicon growth during the necking stage, with $R = 1$ cm, $V_p = 100$ cm/hr and $\sigma_0 = 1$ MPa, the corresponding Dislocation Damköhler Number is $Da = 10$. For InP growth during the necking stage, with $R = 1$ cm, $V_p = 100$ cm/hr and $\sigma_0 = 1$ MPa, $Da = 250$. In the case of silicon for $V_p > 100$ cm/hr, the rate of dislocation multiplication becomes comparable with the motion of the crystal and the dislocation density must be calculated by solution of the coupled equation set (3.67) and (3.68). This is the situation during the growth of silicon sheets studied by Dillon et al. (1987) and Lambropoulos et al. (1983).

Theoretically, if the growth rate can be made extremely large, $Da$ will be small and dislocations will not have time to form. Völkl and Müller have observed this limit in simulations for InP growth by the LEC method. Unfortunately, the coupling between the crystal growth rate and temperature field caused by latent heat release and the growth rate limit for morphological stability of the melt/crystal interface appear to make the high growth rate limit impractical for InP and GaAs growth.

3.5. Summary and Discussion

Analysis of the model of dislocation dynamics developed in Section 3.2 sheds new light on the problems associated with growing low dislocation density materials by the Czochralski and the liquid encapsulated Czochralski methods. The analysis gives a picture of the mechanism for dislocation formation that is relatively independent of the details of the temperature and stress fields, but that distinguishes qualitatively between thermal environments that lead to either increasing or decreasing stress in the crystal as the distance from the interface is increased. This picture is summarized below.
For typical operating conditions, plastic stress relief by dislocation motion is the dominant mechanism for relieving the large thermal stresses induced in the crystal. For the crystalline materials (InP, GaAs, Si) at temperatures near the melting point and at stresses near the CRSS, dislocation formation is so rapid that the dislocation density builds up from the value grown in at the interface to a bulk value in a thin boundary layer of thickness $O(R Da^{-1})$, where the Dislocation Damkohler Number

$$Da = KB_0 \left( \frac{R}{V_p} \right) \exp \left( -\frac{Q}{kT_m} \right) \sigma^{m+1}$$

measures the rate of formation of dislocations relative to the rate of motion of the crystal through the temperature and stress fields.

The details of this boundary layer are unimportant compared to the value of the dislocation density in the bulk crystal, which is created so that the plastic stress relief due to dislocation motion balances the applied stress to the point that no more dislocations form. The analysis predicts expressions for these bulk values that depend on the shape of the axial stress profile in the crystal, as measured by the von Mises stress $\sigma$, a scalar invariant of the elastic stress tensor. For crystal growth conditions, where the stress is either constant or a decreasing function of the axial distance, the bulk dislocation density scales as $\sigma^2$ and is given by eq. (3.60). This is the case for CZ growth of silicon or for LEC growth with very small encapsulant layers.

In LEC systems, where the depth of the encapsulant causes the stress to increasing with axial distance, the dimensional bulk dislocation density scales as $\delta \sigma(z)^2$, where $\delta$ is a dimensionless measure of the back stress created by the dislocations and $\sigma(z)$ is the axial von Mises stress profile; see eq. (3.76). The dislocation density increases with axial distance to a maximum that is determined by the highest value of the von Mises stress in the crystal. The predictions from this analysis are in close agreement with the numerical simulations of Vökl and Müller (1989) using the model of Alexander and Haasen for studies of InP growth. This comparison, coupled with the direct comparison by Vökl and Müller to etch-pit measurements establishes the ability of the model of Alexander and Haasen to quantitatively model dislocation formation.

126
The model of dislocation dynamics does not accurately predict the formation of dislocations in silicon grown by the Czochralski method, where dislocation free material has been routinely produced. An explanation is offered. This explanation is that the number of grown in dislocations at the interface after the Dash necking procedure may be too low to allow effective dislocation multiplication. Then the model of Alexander and Haasen may be useful in explaining why the dislocations disappear during necking. Here, high growth rates and low thermal stresses lead to conditions where the Dislocation Damköhler Number is not large and dislocation multiplication is too slow to be effective. This hypothesis suggests that understanding the breakdown of the growth of dislocation free silicon will hinge on understanding the mechanisms for dislocation generation from microdefects.

The expressions for the bulk dislocation density give simple relationships that show the importance of maintaining low stresses in the crystal by tailoring of the thermal field in crystal growth systems. Only by careful engineering can the temperature gradients and the resulting stresses be kept low enough that the final dislocation density will be low. These expressions serve the same purpose as the critical resolved shear stress condition introduced by Jordan et al. (1980), but result from an entirely different picture of the solid mechanics in the crystal.

Analysis of dislocation multiplication based on the model of Alexander and Haasen (1968) provides a clear picture for the dynamics of dislocations in semiconductor crystals during bulk melt growth and determines regions in the crystal where the dominant mechanisms of defect interactions and dynamics are different. This division of the growing crystal into regions according to the dominant mechanisms of defect dynamics is presented in Figure 3.11. In the melt/crystal interfacial region \(-d < z < 0\) generation of mobile dislocations from intrinsic point defects at concentrations higher than the equilibrium levels is the dominant mechanism and determines the initial condition for the analysis of dislocation dynamics in the bulk. A thin boundary layer of dimensionless thickness \(O(Da^{-1})\) follows where dislocation multiplication dominates. In this layer the dislocation sources are other mobile dislocations and the dislocation density increases abruptly as a result of cross-slip. In this region dislocations move at high speeds because of the high levels of temperature and effective stress. Finally, in the bulk away from the dislocation
Figure 3.11. Decomposition of the growing crystal into three regions according to the dominant mechanisms of defect dynamics. Dislocation generation dominates in the melt/crystal interfacial region, dislocation multiplication dominates in the dislocation density boundary layer and impurity gettering to dislocation cores dominates in the bulk away from the boundary layer.
density boundary layer the dislocation population has reached an equilibrium level and dislocations move at low velocities or dislocation motion is even interrupted in regions where the effective stress becomes zero. In this region the interaction between dislocations and impurities form atmospheres around dislocations or getter inside dislocation cores with further effects on dislocation dynamics. The interaction between dislocations and impurities is the origin of the reduction of dislocation densities from the introduction of dopants in the crystal. This interaction is analyzed in detail for the system of interstitial oxygen in silicon in Chapters 4 and 5.
4. TRANSPORT OF INTERSTITIAL OXYGEN AROUND DISLOCATIONS IN SILICON

The effects of interstitial oxygen in silicon are analyzed as a case study of an impurity in a semiconductor crystal which affects dislocation dynamics with respect to dislocation motion in the intrinsic semiconductor. As emphasized in chapters 2 and 3, impurities play important roles in modifying the mechanical properties of the material during processing. Impurities also reduce dislocation densities during crystal growth because the dislocation multiplication rate is directly related to dislocation mobility. The analysis in this section focuses on the microscopic and macroscopic aspects of the formation of oxygen clouds by oxygen migration around dislocations. The effects of oxygen migration on dislocation dynamics and the mechanical behavior of silicon is analyzed in Chapter 5.

Oxygen is the most important impurity species in silicon grown by the Czochralski method. The main source for the introduction of oxygen in the silicon lattice is the dissolution of the quartz crucibles used in the growth process (Mikkelsen, 1986; Lin and Benson, 1987). Oxygen is an interstitial and electrically inactive impurity in silicon (Mikkelsen, 1986). Czochralski grown silicon crystals typically have oxygen concentrations $c_0$ between $1 \times 10^{17}$ – $1 \times 10^{18}$ atoms/cm$^3$, while in high-purity silicon grown by the floating zone method (Brown, 1988) the oxygen level is below $1 \times 10^{16}$ atoms/cm$^3$ (Sumino et al., 1980). The increased mechanical strength of silicon crystals grown by the Czochralski method with respect to high-purity silicon is of major importance in device processing, where the warpage of silicon wafers during thermal cycling must be prevented (Sumino, 1987a).

Oxygen transport is analyzed around the mobile dislocations, which control the mechanical behavior of the crystal. Calculation of the oxygen distribution around mobile dislocations is the basis for the analysis of the effects of oxygen on dislocation motion and consequently on the mechanical strength of silicon. The starting point of the analysis is the development of a systematic model of interstitial impurity migration under the action of a force field. This model is presented in Section 4.1, where constitutive equations are developed for the drift velocity and the diffusivity tensor of interstitial impurities, as functions of the applied force field and temperature. The theory of interstitial impurity migration is
used in Section 4.2 to derive a constitutive model for the anisotropic migration of oxygen in the stress field created by 60° dislocations in silicon. The oxygen distribution around immobile and gliding dislocations in silicon is calculated in Section 4.3, where the formation of an oxygen atmosphere around the dislocation core is modeled and the evolution of the atmosphere with increasing dislocation velocity is computed.

4.1. Constitutive Modeling of Interstitial Impurity Diffusion and Drift

Quantitative modeling of impurity migration in the presence of dislocations and the subsequent analysis of the effect of the impurity on the dislocation motion requires analysis of the effect of the stress field on the migration of the impurity. The first step in our analysis is the development of a theory for point defect and impurity migration in cubic-type lattices in the presence of an internal force established by the interaction of the migrating interstitial particle and a dislocation. A constitutive model for interstitial impurity migration is developed as a result of the theory.

The motion of point defects, composed of either native lattice defects or impurities, has been studied in great detail (Flynn, 1972; Girifalco, 1973), but little is known about the effect of an applied or internal stress field on defect transport (Girifalco and Grimes, 1961; Koehler, 1969; Viñals and Sekerka, 1988). The complexity of incorporating the stress field is best understood by picturing the migration of a light interstitial impurity atom as a jumping process between interstitial sites in a cubic lattice. The interaction between the impurity atom and the stress field caused by a distant dislocation biases the probability of the atom jumping to neighboring sites and leads to anisotropic migration that is a function of the direction and magnitude of the interaction force. Although molecular dynamics and Monte Carlo calculations have been used to demonstrate many of the atomic-level features of particle migration in the vicinity of large-scale crystalline defects (Kwok et al., 1984; Lancon et al., 1985; Kirchheim, 1987; Kirchheim and Stolz, 1988; Limoge and Bocquet, 1988), no unifying theory has emerged that links these calculations to the concepts of diffusivity and drift inherent in a continuum theory for the particle flux.

A systematic theory is presented in this section that links atomistic and continuum
scales for the migration of particles in cubic-type lattices acted on by a force caused by an applied or internal stress field. The development is based on the connections between ensemble averaging starting from the stochastic analysis of jump rate theory and a macroscale transport theory for particle density including both particle diffusion and drift due to the applied force. This link is established using the formal method of moments described by Brenner (1980) and yields expressions for the \textit{continuum} diffusivity tensor and drift velocity in terms of the first and second moment tensors of the point defect number density distribution computed on the \textit{atomistic} scale. The macroscopic properties depend on the stress field and the resulting interaction force through the altering of the activation barrier height for particle jumps and the biasing of the jump frequencies in different directions. The outline of the methodology that is followed for the theoretical analysis of interstitial impurity migration is presented in Figure 4.1.

The jump rate theory for point defect migration is systematically reviewed in Section 4.1.1. Monte Carlo methods are introduced for the simulation of the impurity jump process in Section 4.1.2. The macroscopic analysis for particle migration is presented in Section 4.1.3, where expressions for the drift velocity and the diffusivity tensor of the point defects are derived as functions of the first and second moments of the dilute impurity distribution function. The results of the moment analysis and stochastic modeling of the atomic jump process in a two-dimensional square lattice are used in Section 4.1.4 to derive a closed-form solution for the dependence on a constant interaction force of the drift velocity and the diffusivity tensor for an interstitial impurity. Interstitial diffusion under constant force also is analyzed by Monte Carlo simulations to reproduce particle trajectories in the two-dimensional square lattice.

\textbf{4.1.1. Jump Rate Theory for Point Defect Migration}

The statistical mechanics of point defect migration has been studied in detail (Flynn, 1972; Girifalco, 1973). The migration of an isolated point defect on a lattice is modeled as a thermally activated jump process between nearest neighbor equilibrium lattice sites (Kehr and Binder, 1984). If the migrating particle, either substitutional or interstitial,
Figure 4.1. Outline of the methodology for the theoretical analysis of interstitial impurity migration.
occupies an equilibrium site $i$, it can move from this site to $Z$ nearest neighbor sites, where $Z$ is the coordination number of the lattice. The path from the $i$th to the $j$th site is characterized by a jump frequency $w_{ij}$. Using the formulation of the rate process theory in solids introduced by Vineyard (1957), the jump frequency $w_{ij}$ is given by

$$w_{ij} = w_0 \exp(-Q_{ij}/kT), \quad (4.1)$$

where $Q_{ij}$ is the height of the activation energy barrier for the $i \rightarrow j$ jump, $k$ is Boltzmann's constant and $w_0$ is the attempt frequency, which is independent of the energy fluctuations (Limoge and Bocquet, 1988). The equivalence of eq. (4.1) with an analogous equation that involves the activation Gibbs free energy for the $i \rightarrow j$ jump, $\Delta G_{ij}$, has been discussed by Girifalco (1973). The term $\exp(-\Delta G_{ij}/kT)$ in the latter equation is derived from eq. (4.1) using the high-temperature approximation for the entropy of a harmonic system. This equivalence is analyzed in Chapter 6 for the migration of intrinsic point defects. The description of the $i \rightarrow j$ jump as a Poisson process was given by Lancon et al. (1985) and Limoge and Bocquet (1988) for the case of random walk interstitial diffusion in periodic disordered media.

In the case of random walk diffusion in a perfect lattice, the activation energy barriers are the same for all the sites, $Q_{ij} = Q^\circ$, and eq. (4.1) takes the simple form

$$w_{ij} = w_0 \exp(-Q^\circ/kT). \quad (4.2)$$

However, in the presence of an applied or internal strain field, the activation barrier changes because of two effects (Koehler, 1969). The first is the drift effect associated with changes in the equilibrium point defect energy with position due to the interaction with the stress field created in the lattice by the point defect. The second is the barrier effect which results because stress fields alter the height of the activation barrier (Girifalco and Grimes, 1961; Viñals and Sekerka, 1988).

Calculation of the activation energy barrier $Q_{ij}$ in a strained crystal requires systematic analysis based on rate process (transition state) theory in solids (Vineyard, 1957). According to the theory the jump frequency $w_{ij}$ is given by

$$w_{ij} = \left(\frac{kT}{2\pi m}\right)^{1/2} \frac{\int_{A_s} dA_s \exp(-\varphi/kT)}{\int_A \exp(-\varphi/kT) dA}. \quad (4.3)$$
where the integrals are calculated over the regions $A_s$ and $A$ of the configurational space, respectively. The region $A_s$ is a hypersurface in the configuration space defined such that the surface passes through the point corresponding to the migrating atom at its activated position with all other atoms at their equilibrium positions. The hypersurface $A_s$ is perpendicular to contours of constant potential energy and divides the configuration space into two symmetric parts. The configuration volume $A$ is one of these parts. The factor that multiplies the ratio of the two configurational integrals is the result of the ratio of the integrals over the momenta of the exponent of the kinetic energy part of the system Hamiltonian. The factor $(kT/2\pi m)^{1/2}$ is rigorous only in the case of vacancy and self-interstitial migration. The differences in the masses of an impurity and host atoms changes the above result. However, we are interested only in the configurational integrals to calculate the ratio $w_{ij}(\xi)/w_{ij}(0)$ of the jump frequencies in the strained and unstrained lattice, respectively where $\xi(r)$ is the strain field. Viñals and Sekerka (1988) calculated the configurational integrals in (4.3) using a Monte Carlo technique for a Lennard-Jones solid and discussed the assumptions of transition state theory, where particle jumps are treated as uncorrelated equilibrium fluctuations.

In the strained lattice the interatomic potential is a function of all the atomic coordinates $\{q_i\}$ and the six independent components of the symmetric strain tensor $\xi$, i.e. $\varphi = \varphi(q_i, \xi)$. The potential $\varphi$ is expanded in a Taylor series in the strains about the point of zero strain. When the migrating particle is at the equilibrium position characterized by the position vector $r_e$ the expansion yields

$$\varphi = \varphi(q_i, 0) + U_e(\xi) + C^{(1)} : \xi + C^{(2)} :: \xi \xi + O(\epsilon^3), \quad (4.4)$$

where $U_e(\xi)$ is the interaction potential between the strain field and the migrating point defect at the equilibrium point. The second rank $C^{(1)}$ and the fourth rank $C^{(2)}$ tensors are defined in component form by

$$C^{(1)}_{\alpha \beta} \equiv \left( \frac{\partial^2 \varphi}{\partial \epsilon_{\alpha \beta}} \right)_{q_i, 0}, \quad (4.5a)$$

$$C^{(2)}_{\alpha \beta \rho \sigma} \equiv \frac{1}{2} \left( \frac{\partial^2 \varphi}{\partial \epsilon_{\alpha \beta} \partial \epsilon_{\rho \sigma}} \right)_{q_i, 0}, \quad (4.5b)$$

135
where \((\alpha, \beta, \rho, \sigma)\) are cartesian indices and the subscripts outside of the parentheses indicate that the derivatives are evaluated when the strains are zero and the coordinates of the \(N\) (or \(N+1\) in the case of a solute atom) particles have the values \(\{q_j\}, j = 1, \ldots, 3N\) or \(3N+3\). If the migrating point defect is at the saddle point, the expansion for the potential \(\varphi^*\) yields

\[
\varphi^* = \varphi^*(q_i, 0) + U_s(\varepsilon) + C^{(1)*} : \varepsilon + C^{(2)*} : \varepsilon \varepsilon + O(\varepsilon^3),
\]

where the stars denote that the corresponding quantities are calculated with the point defect at the saddle point position characterized by the position vector \(r_s\), and \(U_s\) is the potential of the interaction between the strain field and the point defect at the saddle point.

The statistical average of a quantity \(M\) over a particular region of the configurational space is denoted by angular brackets as

\[
\langle M \rangle_A \equiv \frac{\int_A M \exp(-\varphi(q_i, 0)/kT) \, dA}{\int_A \exp(-\varphi(q_i, 0)/kT) \, dA} ,
\]

\[
\langle M \rangle_{A_s} \equiv \frac{\int_{A_s} M^* \exp(-\varphi^*(q_i, 0)/kT) \, dA_s}{\int_{A_s} \exp(-\varphi^*(q_i, 0)/kT) \, dA_s} .
\]

In eqs. (4.4) and (4.6), \(U_e\) and \(U_s\) do not depend on the unstrained coordinates of the particle; therefore, eqs. (4.3), (4.4), and (4.6) give the jump frequency ratio as

\[
\frac{w_{ij}(\varepsilon)}{w_{ij}(0)} = \exp\left(-\frac{U_s - U_e}{kT}\right) \frac{\langle \exp\left(-\frac{1}{kT} C^{(1)} : \varepsilon\right) \exp\left(-\frac{1}{kT} C^{(2)} : \varepsilon \varepsilon\right) \rangle_{A_s}}{\langle \exp\left(-\frac{1}{kT} C^{(1)} : \varepsilon\right) \exp\left(-\frac{1}{kT} C^{(2)} : \varepsilon \varepsilon\right) \rangle_A} .
\]

The exponents inside the angular brackets are expanded for small strains to give

\[
\frac{\langle \exp\left(-\frac{1}{kT} C^{(1)} : \varepsilon\right) \exp\left(-\frac{1}{kT} C^{(2)} : \varepsilon \varepsilon\right) \rangle_{A_s}}{\langle \exp\left(-\frac{1}{kT} C^{(1)} : \varepsilon\right) \exp\left(-\frac{1}{kT} C^{(2)} : \varepsilon \varepsilon\right) \rangle_A} \approx \frac{1 - \frac{1}{kT} [(C^{(1)})_{A_s} : \varepsilon + (C^{(2)})_{A_s} : \varepsilon \varepsilon + O(\varepsilon^3)]}{1 - \frac{1}{kT} [(C^{(1)})_A : \varepsilon + (C^{(2)})_A : \varepsilon \varepsilon + O(\varepsilon^3)]} .
\]
Taking the logarithms of both sides of eq. (4.8) and keeping only first order terms in

\[ \ln \left[ \frac{w_{ij}(\varepsilon)}{w_{ij}(0)} \right] \approx - \left( \frac{U_s - U_e}{kT} \right) + \frac{1}{kT} \left[ \langle C^{(1)} \rangle_A - \langle C^{(1)} \rangle_{A_s} \right] : \varepsilon + \left[ \langle C^{(2)} \rangle_A - \langle C^{(2)} \rangle_{A_s} \right] : \varepsilon \varepsilon. \]  

(4.10)

The potential energies of the interaction between the strain field and the point defect at the equilibrium and the saddle point positions are \( U_s(\varepsilon) = U_s(r_s) \) and \( U_e(\varepsilon) = U(r_e) \). Therefore, \( U(r_s) \) is expanded in a Taylor series about the equilibrium position of the point defect to give

\[ U_s - U_e = \nabla U \cdot (r_s - r_e) + \frac{1}{2} \nabla \nabla U : (r_s - r_e)(r_s - r_e) + O(r_s - r_e)^3, \]  

(4.11)

Equations (4.10) and (4.11) give

\[ w_{ij}(\varepsilon) = w_{ij}(0) \exp \left\{ - \frac{1}{kT} \left[ \nabla U \cdot (r_s - r_e) + \frac{1}{2} \nabla \nabla U : (r_s - r_e)(r_s - r_e) - \right. \right. \]

\[ \left. \left. \left[ \langle C^{(1)} \rangle_A - \langle C^{(1)} \rangle_{A_s} \right] : \varepsilon - \left( \langle C^{(2)} \rangle_A - \langle C^{(2)} \rangle_{A_s} \right) : \varepsilon \varepsilon \right\} \right\}. \]  

(4.12)

Using eqs. (4.12) and (4.2) for the jump frequency in the unstrained lattice gives eq. (4.1) for the jump frequency in the strained lattice with a height of the activation energy barrier equal to

\[ Q_{ij} \approx Q^0 + \nabla U \cdot (r_s - r_e) + \frac{1}{2} \nabla \nabla U : (r_s - r_e)(r_s - r_e) - \]

\[ \left( \langle C^{(1)} \rangle_A - \langle C^{(1)} \rangle_{A_s} \right) : \varepsilon - \left( \langle C^{(2)} \rangle_A - \langle C^{(2)} \rangle_{A_s} \right) : \varepsilon \varepsilon. \]  

(4.13)

Equation (4.13) is rewritten as

\[ Q_{ij} = Q^0 + \nabla U \cdot (r_s - r_e) + \frac{1}{2} \nabla \nabla U : (r_s - r_e)(r_s - r_e) - S^{(1)} : \varepsilon - S^{(2)} : \varepsilon \varepsilon, \]  

(4.14)

where the tensors \( S^{(1)} \) and \( S^{(2)} \) are of second and fourth order, respectively, and are defined by

\[ S^{(1)} \equiv \langle C^{(1)} \rangle_A - \langle C^{(1)} \rangle_{A_s}, \]  

(4.15a)
\[ S^{(2)} \equiv \langle C^{(2)} \rangle_A - \langle C^{(2)} \rangle_A^* , \]  

(4.15b)

In the Taylor series expansion for the interatomic potential \( \varphi \) that yielded eq. (4.14), terms have been retained up to second order in strain and up to second order in \( r_s - r_e \). Eqs. (4.1) and (4.14) provide the basis for our analysis.

If the interaction potential \( U(\mathbf{r}) \) is not uniform throughout the lattice, a force \( \mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r}) \) acts on the migrating particle and biases the particle jump directions. This force field is generally classified as either homogeneous or nonhomogeneous. An example of a homogeneous force acting on a migrating interstitial is given by the Gorski effect, which occurs during the bending of thin rectangular metallic plates (Kronmüller et al., 1971). Here the strain field created by the bending interacts with a point defect via a linear interaction potential and produces \( \varphi \) constant applied force on the defect. A typical example of a nonhomogeneous interaction force \( \mathbf{F}(\mathbf{r}) \) is provided by the first-order linear elastic interaction between an edge dislocation and a point defect. In this case the interaction potential is \( U(\mathbf{r}) = K \sin \theta/r \), where the dislocation lies along the \( z \)-axis of a cylindrical coordinate system, and \( K \) depends on the elastic constants of the crystal and the Burgers vector of the dislocation (Hirth and Lothe, 1982; Nabarro, 1987).

Comparing the drift and the barrier terms in eq. (4.14) leads to simplified expressions for \( Q_{ij} \) for certain diffusion mechanisms. Koehler (1969) proved that in the case of vacancy diffusion in the stress field created by defects such as interstitial clusters and edge dislocations, the first order barrier terms dominate over the second order barrier terms for distances higher than a few lattice parameters away from the defect, justifying the truncation of the Taylor series. This truncation also is valid for the migration of self-interstitials and substitutional solute atoms. It is proved in Section 4.1.5 that the first order drift terms are dominant over the second order drift terms for point defect migration in the stress field of an edge dislocation. Therefore, second order effects can be neglected in the study of the migration of point defects and eq. (4.14) is simplified to

\[ Q_{ij} \simeq Q^0 + \nabla U \cdot (r_s - r_e) + S^{(1)} : \xi. \]  

(4.16)

Girifalco and Grimes (1961) developed the relationship between \( S^{(1)} \) and the activation volume for migration \( \Delta V^* \). For the case of uniform compression or expansion, \( S^{(1)} = S^{(1)} \mathbf{I} \),
the relationship is
\[
\Delta V^* \equiv \left[ \frac{\partial (\Delta G)}{\partial P} \right]_T = S^{(1)} \beta,
\]
where \( I \) is the identity tensor, \( \beta \) is the compressibility of the material and \( \Delta G \) is the Gibbs free energy change for the formation of the activated state configuration. The change in the Gibbs free energy for the formation of native lattice defects is included in \( \Delta G \) if such point defects are involved in the diffusion mechanism. Eq. (4.17) implies that very low positive values of \( \Delta V^* \) imply negligible barrier effects for the migration process. Birnbaum et al. (1971) studied the effects of diffusivity gradients caused by pressure on point defect diffusion around dislocations and reviewed the theoretical and experimental results for activation volumes of migration. Their most important conclusion was that for light interstitial solute atoms \( \Delta V^* \ll \Delta V^f \), where \( \Delta V^f \) is the formation volume due to the size misfit of the solute atom; \( \Delta V^f \) is proportional to \( U \). Therefore, for the migration of interstitial impurities eq. (4.16) is simplified to
\[
Q_{ij} \simeq Q^o + \nabla U \cdot (r_s - r_e).
\]

We focus attention on interstitial impurity migration in a stress field which interacts with the impurity via the potential field \( U(r) \), the gradient of which defines the interaction force \( F(r) \). We assume that the saddle point position for a jump from a site \( i \) to a nearest neighbor site \( j \) is located in the middle of the distance between the two sites. This assumption has been carefully discussed by Viñals and Sekerka (1988). If \( l \) is the distance between nearest neighbor sites eq. (4.18) gives
\[
Q_{ij} \simeq Q^o - \frac{1}{2} l F(r_e) \cdot \hat{r}_{ij},
\]
where \( \hat{r}_{ij} \) is the unit vector in the \( i \rightarrow j \) direction. One-dimensional representations are shown in Fig. 4.2 for the potential energy variation with respect to the lattice positions for the cases of zero, homogeneous and inhomogeneous forces \( F \), respectively. The jump frequency \( \omega_{ij} \) is given generally by an expression of the form (4.1).

According to the above description, the action of the force \( F \) on a migrating interstitial solute is the dominant mechanism that creates preferred jump directions. Therefore, a
rigorous study of the migration process requires the calculation of the elements of the anisotropic and spatially dependent diffusivity tensor \( D \) of the defect in the lattice rather than a scalar diffusion coefficient. The development of a systematic procedure for the calculation of \( D \) requires the explicit dependence of the diffusivity tensor on the force \( F \) under constant temperature, i.e. \( D = D(F(r)) \). If \( F = 0 \), the diffusivity tensor in a cubic-type lattice is isotropic, \( D = DI \), where \( I \) is the identity tensor. The presence of \( F \) lowers the symmetry of the crystal creating preferred directions. In addition, the force \( F \) introduces a drift mechanism for point defect migration. Random walk diffusion models are not sufficient, and the drift velocity \( u_d \) of the point defect due to the action of \( F \) must be calculated, i.e. \( u_d = u_d(F(r)) \).

4.1.2. Monte Carlo Methods for Simulation of Interstitial Impurity Migration

Monte Carlo methods are well suited to the study of diffusion in solids (Kehr and Binder, 1984; Lancon et al., 1985; Kirchheim, 1987; Kirchheim and Stolz, 1988; Limoge and Bocquet, 1988) because the steps in the Monte Carlo simulation can be directly related to real time for the description of the microscopic kinetics of the diffusion process. For such kinetic processes the Monte Carlo simulation is a method for solving the master equation with the detailed balance condition satisfied at thermal equilibrium (Binder, 1986; Van Kampen, 1987; Kang and Weinberg, 1989) and based on the expression for the transition probability \( W(R \rightarrow R') \) that the system changes state from \( R \) to \( R' \) within a time unit. For a thermally activated process the appropriate probability expression is

\[
W(R \rightarrow R') = W_{ij} = \exp(-Q_{ij}/kT),
\]

(4.20)

where \( Q_{ij} \) is the activation energy. Kang and Weinberg (1989) discussed the advantages of the probability expression (4.20) compared to other expressions that satisfy the detailed balance condition.

The probability \( W_{ij} \) has been used for the study of interstitial diffusion in solids (Kirchheim, 1987; Limoge and Bocquet, 1988), where the correct time unit for the simulated time is the Monte Carlo step per particle. For a single diffusing particle, the Monte Carlo step is equal to the time for the particle to attempt one jump, which is equivalent
Figure 4.2. Diagram of one-dimensional potential energy $\varphi$ with respect to the lattice positions $r$ for the cases of (a) zero, (b) constant and (c) inhomogeneous interaction force $F(r)$. 
to choosing
\[ w_0 = \frac{1}{Z}. \] (4.21)
in eq. (4.1).

We use the Monte Carlo (MC) simulation to reproduce particle trajectories for interstitial diffusion in cubic-type lattices under the influence of force fields. The algorithms that are been used are modifications of methods by Lançon et al. (1985) and Limoge and Bocquet (1988). The generation of each trajectory consists of two steps. First, the energy fluctuation \( E \) associated with the step is calculated using a pseudo-random number \( RND \) in the interval \([0,1]\)
\[ RND \equiv \exp(-E/kT). \] (4.22)
The particle jumps if \( E \) is higher than the energy barrier in eq. (4.20). If \( F = 0, \) \( E > Q^o \) is the criterion for a successful jump; if \( F \neq 0, \) the particle jumps from the site \( i \) if \( E \) is higher than the lowest energy barrier \( \min_j \{Q_{ij}\} \) of the surrounding lattice sites, where the \( \{Q_{ij}\} \) are calculated according to eq. (4.19).

Second, the jump direction is chosen among all the \( Z \) nearest neighbors. If \( F = 0 \) the choice is made by random sampling. If \( F \neq 0, \) the preferred directions are taken into account using a pseudo-random number \( RND \) in the interval \([0,1]\) and the criterion for the choice of the jump direction:
\[ \sum_{j=1}^{l-1} w_{ij}/w_i < RND < \sum_{j=1}^{l} w_{ij}/w_i, \] (4.23)
where \( l \) is the index for the position where the particle is going to jump and \( w_i \equiv \sum_{j=1}^{Z} v_{ij}. \)
4.1.3. Macroscopic Analysis of Point Defect Migration

The results of the Monte Carlo simulation are connected to macroscopic diffusion by continuum transport theory. The theory presented in this Section is valid for both interstitial and substitutional diffusion in any lattice structure, but is restricted to the limit of a dilute solute in the crystal, so that the dependences of the diffusivity and drift velocity on concentration can be ignored. This condition is met by oxygen in Czochralski growth, where the typical number density is $10^{18}$ cm$^{-3}$. The dilute asymptotic limit also is equivalent to assuming that the diffusing particles do not interact. The strain fields created by point defects in the lattice interact by weak short-range interactions; these are proportional to $r^{-6}$, where $r$ is the distance between individual defects (Eshelby, 1976). The results of the theory are applied to the case of interstitial impurity migration. The assumption of the transport of dilute intrinsic point defects and impurities and impurities is appropriate for point defects in metals and semiconductors under equilibrium conditions because of the high formation energies of point defects in these materials (Balluffi and Granato, 1979; Chikawa, 1987). The same assumption is valid for dopants in semiconductor crystals where dopant levels are typically much lower than $10^{-3}$ atomic fraction (Ghandi, 1983; Sumino, 1987a).

4.1.3.1. Transport Equation for the Number Density of Point Defects

The development begins with the probability density $P(r,t)$ that the displacement vector of an isolated point defect from its initial position is at position $r$ at time $t$. The flux vector $\mathbf{J}$ for the probability is given by Fick's law as

$$ \mathbf{J} = u_d P - \mathbf{D} \cdot \nabla P, \quad (4.24) $$

where $u_d$ is the drift velocity and $\mathbf{D}$ is the diffusivity tensor.

Conservation of the probability density requires that $P$ satisfy the continuity equation

$$ \frac{\partial P}{\partial t} = -\nabla \cdot \mathbf{J}. \quad (4.25) $$

143
Combining eqs. (15) and (16) gives the continuity equation for the probability $P$

$$\frac{\partial P}{\partial t} = \nabla \cdot (u_d P) + \nabla \cdot (D \cdot \nabla P),$$  

(4.26)

which also is the Fokker-Planck equation for a particle diffusing in a three-dimensional inhomogeneous and anisotropic medium with a drift velocity $u_d$ (Van Kampen, 1987).

The distribution of point defects is characterized by the number density $n(r,t)$. Because only defect migration occurs, the total number of point defects $N_{pd}$ in the infinite lattice is conserved and is written as

$$N_{pd} = \int_V n(r,t)d^3r.$$  

(4.27)

The number density $n(r,t)$ is expressed in terms of the initial value $n(r,0)$ and the solution $P(r,t)$ of eq. (4.26) under the appropriate boundary conditions. Using the linear superposition principle (Brenner, 1980) gives

$$n(r,t) = \int_V n(r',0)P(r,t|r')d^3r'.$$  

(4.28)

The transport equation for the number density $n(r,t)$ is obtained from eq. (4.28) and the linearity of eq. (4.26) as

$$\frac{\partial n}{\partial t} = -\nabla \cdot (u_d n) + \nabla \cdot (D \cdot \nabla n).$$  

(4.29)

The appropriate boundary condition for the probability $P$ defined on an infinite lattice is

$$\lim_{|r|\to\infty} P(r,t) = 0.$$  

(4.30)

In addition, $P$ must vanish with a sufficiently rapid rate with distance to insure convergence of the integral in the normalization condition $\int_V Pd^3r \equiv 1$. Therefore, eqs. (4.28) and (4.30) give the appropriate boundary condition for $n(r,t)$ as

$$\lim_{|r|\to\infty} n(r,t) = 0,$$  

(4.31)
where \( n \) vanishes with a sufficiently rapid rate as \( |r| \to \infty \) for the integral (4.27) to converge. Equations (4.29) and (4.31) define a boundary-value problem for the number density \( n(r, t) \) which can be solved if \( u_d \) and \( D \) are known functions of \( r \). Equation (4.29) becomes the starting point of our analysis and its solution provides the basis for the quantitative modelling of the dynamics of point defects in the crystal.

4.1.3.2. Moment Analysis of the Transport Equation

The method of moments is used to derive expressions for \( u_d \) and \( D \) as functions of the moments of the probability density \( P(r, t) \). This method has been used for the study of the Brownian motion of a spherical tracer particle in a fluid that moves in spatially periodic flows (Brenner, 1980) and for the Monte Carlo simulation of charge transport in semiconductors (Price, 1979; Jacoboni and Reggiani, 1983). We extend the approach of Jacoboni and Reggiani (1983) to a three-dimensional medium.

The \( m \)-th moments of the number density are polyadics defined as

\[
M_m \equiv \frac{1}{N_{pd}} \int_V r^m n(r, t) d^3r, \tag{4.32}
\]

and are equivalent to the definitions of the ensemble averages of the displacement vector \( r \), the dyadic \( rr \) and so on, in a subensemble that is characterized by a given probability distribution \( p(r_0) \). The averages in this subensemble are equilibrium averages in the limit \( t \to \infty \) and are given by

\[
\langle r^m \rangle \equiv \int_V r^m P(r, t) d^3r. \tag{4.33}
\]

Equations (4.27) and (4.28) also imply that

\[
\langle r^m \rangle = M_m. \tag{4.34}
\]

Expressions for \( u_d \) and \( D \) are derived taking the time derivatives of the first and second moments \( M_1 \) and \( M_2 \). To do this, eq. (4.29) is rewritten as

\[
\frac{\partial n}{\partial t} = -\nabla \cdot J_1 + \nabla \cdot J_2, \tag{4.35a}
\]
where the two fluxes $J_1$ and $J_2$ are defined as

$$J_1 \equiv u_d n,$$  \hspace{1cm} (4.35b)

$$J_2 \equiv D \cdot \nabla n.$$  \hspace{1cm} (4.35c)

Differentiating the expression for $M_1$ with respect to $t$ and using eq. (4.35a) gives

$$\frac{dM_1}{dt} = \frac{1}{N_{pd}} \int_V r \frac{\partial n}{\partial t} d^3r = \frac{1}{N_{pd}} \left\{ - \int_V r(\nabla \cdot J_1) d^3r + \int_V r(\nabla \cdot J_2) d^3r \right\}. \hspace{1cm} (4.36)$$

Using the divergence theorem and the boundary condition (4.31) reduces the first integral in (4.36) to

$$\int_V r(\nabla \cdot J_1) d^3r = \int_V J_1 d^3r. \hspace{1cm} (4.37a)$$

The second integral in eq. (4.36) is simplified as

$$\int_V r(\nabla \cdot J_2) d^3r = \oint_{\partial V} dS \cdot [rJ_2]_{\partial V} + \int_V J_2 d^3r, \hspace{1cm} (4.37b)$$

where $\partial V$ denotes the boundary of the domain ($V$) and $S$ is the boundary area vector. The integrand $[rJ_2]$ vanishes on the boundary $\partial V$ because $n$ is identically zero and $\nabla n$ vanishes due to the rapid decay rate of $n$. The time derivative of the first moment (4.36) reduces to

$$\frac{dM_1}{dt} = \frac{1}{N_{pd}} \left\{ \int_V J_1 d^3r - \int_V J_2 d^3r \right\}. \hspace{1cm} (4.38)$$

Taking the time derivative of the second moment and using eq. (4.35a) gives

$$\frac{dM_2}{dt} = \frac{1}{N_{pd}} \int_V r \frac{\partial n}{\partial t} d^3r = \frac{1}{N_{pd}} \left\{ - \int_V rr(\nabla \cdot J_1) d^3r + \int_V rr(\nabla \cdot J_2) d^3r \right\}. \hspace{1cm} (4.39)$$

Using the divergence theorem, the boundary condition (4.31) and eq. (4.27) gives

$$\frac{dM_2}{dt} = \frac{1}{N_{pd}} \left\{ \int_V J_1 rd^3r + \int_V [J_1 r]^T d^3r - \int_V J_2 rd^3r - \int_V [J_2 r]^T d^3r \right\}, \hspace{1cm} (4.40)$$

where $\cdot^T$ denotes the transpose of a matrix.
Following the results of Section 4.1.1 for interstitial migration, if there is no force acting on the migrating particles, the drift velocity is equal to zero and $\mathbf{J}_1 = 0$. Moreover, $\mathbf{D}$ is spatially homogeneous in such a case. Using the divergence theorem, the boundary condition (4.31) and the normalization condition (4.27) reduces the remaining integrals in (4.38) to

\[
\int_V \mathbf{J}_2 d^3r = \mathbf{D} \cdot \int_V (\nabla n) d^3r = 0, \quad (4.41a)
\]

\[
\int_V \mathbf{J}_2 r d^3r = \int_V \mathbf{D} \cdot (\nabla n) r d^3r = -\mathbf{D} N_{pd}. \quad (4.41b)
\]

Combining eqs. (4.41) with eqs. (4.38) and (4.40) gives

\[
\frac{d(r)}{dt} = 0, \quad (4.42)
\]

\[
\frac{d(r)}{dt} = \mathbf{D} + \mathbf{D}^T = 2\mathbf{D}, \quad (4.43)
\]

where Onsager's relations (de Groot and Mazur, 1984) have been used to guarantee that $\mathbf{D}$ is symmetric. The diffusion tensor defined by eq. (4.43) is not necessarily isotropic. No statement on the isotropy of $\mathbf{D}$ can be derived from the moment analysis, because the lattice structure has not been incorporated into the result.

If the force $\mathbf{F}$ that acts on the migrating interstitial particles is uniform throughout the lattice, $\mathbf{u}_d$ and $\mathbf{D}$, which according to the results of Section 4.1.1 are functions of $\mathbf{F}$ alone, are not spatially dependent. In this case, the integrals in eqs. (4.38) and (4.40) reduce to

\[
\int_V \mathbf{J}_1 r d^3r = \mathbf{u}_d \int_V n d^3r = \mathbf{u}_d N_{pd}, \quad (4.44a)
\]

\[
\int_V \mathbf{J}_1 r d^3r = \mathbf{u}_d \int_V n r d^3r = \mathbf{u}_d M_1 N_{pd}, \quad (4.44b)
\]

\[
\int_V \mathbf{J}_2 d^3r = \mathbf{D} \cdot \int_V (\nabla n) d^3r = 0, \quad (4.44c)
\]

\[
\int_V \mathbf{J}_2 r d^3r = \int_V \mathbf{D} \cdot (\nabla n) r d^3r = -\mathbf{D} N_{pd}. \quad (4.44d)
\]

147
and the equation for the first moment (4.38) reduces to

\[ \frac{d(r)}{dt} = u_d. \quad (4.45) \]

The result of the second moment (4.40) is

\[ \frac{dM_2}{dt} = \frac{d}{dt} (M_1 M_1) + 2D, \quad (4.46a) \]

or equivalently

\[ D = \frac{1}{2} \frac{d}{dt} ((rr) - (r)(r)), \quad (4.46b) \]

so that the diffusivity tensor \( D \) is related explicitly to the second central moment of the point defect number density distribution.

In the case where the force field \( F(r) \) is inhomogeneous, moment analysis does not give explicit relations between \( u_d \) and \( D \) and the moments \( M_1 \) and \( M_2 \). The problem of diffusion in an inhomogeneous force field is addressed in Section 4.1.5, as an extension of the problem for a homogeneous force.

4.1.4. Impurity Migration Under Constant Force

4.1.4.1. Analytical Results

The results of the moment analysis are used to derive a constitutive model expressed in terms of the drift velocity \( u_d = u_d(F, T) \) and the diffusivity tensor \( D = D(F, T) \) for the migration of interstitial defects in a two-dimensional square lattice at a given temperature \( T \). The external force \( F \) is assumed to be uniform throughout a square lattice described using the cartesian axes \( x \) and \( y \) to represent the [10] and [01] axes. The force is written as \( F = F_x \hat{x} + F_y \hat{y} \), where \( \hat{x} \) and \( \hat{y} \) are the corresponding unit vectors. For a square lattice, the coordination number is \( Z = 4 \) and the index \( j \) is used to represent the four directions of jump.

Following the analysis of Section 4.1.1, the jump frequencies \( w_{ij} \) are

\[ w_{ij} = C \exp \left( -\frac{Q^o}{kT} \right) \exp \left( \frac{F_j a}{2kT} \right), \quad i = 1, \ldots, 4, \quad j = 1, \ldots, 4, \quad (4.47) \]
where $F_j$ is a component of the set \{\(F_x, -F_x, F_y, -F_y\)\} and $C$ is a frequency. The total jump frequency is

\[
\begin{aligned}
w \equiv \sum_{j=1}^{Z} w_{ij} &= 2C \exp\left(-\frac{Q^o}{kT}\right) \left[ \cosh\left(\frac{F_x a}{2kT}\right) + \cosh\left(\frac{F_y a}{2kT}\right) \right],
\end{aligned}
\tag{4.48}
\]

and the probability $p_j$ that a jump occurs in the $j$ direction is

\[
\begin{aligned}
p_j \equiv \frac{w_{ij}}{w} &= \frac{\exp\left(\frac{F_j a}{2kT}\right)}{2\left[ \cosh\left(\frac{F_x a}{2kT}\right) + \cosh\left(\frac{F_y a}{2kT}\right) \right]}.
\end{aligned}
\tag{4.49}
\]

The total jump frequency is calculated as

\[
w = Z w_0 P_{\text{jump}},
\tag{4.50}
\]

where $P_{\text{jump}}$ is the probability that an attempted jump is successful. This probability is calculated from the lowest energy barrier that the migrating particle has to overcome, i.e.

\[
P_{\text{jump}} = \exp\left[-\frac{Q^o - \max\{F_x, F_y\} a}{kT}\right],
\tag{4.51}
\]

so that the jump frequency $w$ is

\[
w = 4w_0 \exp\left(-\frac{Q^o}{kT}\right) \exp\left[\frac{\max\{F_x, F_y\} a}{2kT}\right],
\tag{4.52}
\]

In the case $F = 0$, eq. (4.52) reduces to the well known result of random walk diffusion on a square lattice (Manning, 1968). Equation (4.52) also implies that the frequency $C$ in eq. (4.47) is

\[
C = 2 \exp\left[\frac{\max\{F_x, F_y\} a}{2kT}\right] \exp\left(\frac{F_x a}{2kT}\right) + \cosh\left(\frac{F_y a}{2kT}\right) - w_0,
\tag{4.53}
\]

and can be interpreted as an enhanced attempt frequency, due to the increased amplitude of the oscillations of the particle at an equilibrium site caused by the applied force; if $F = 0$, $C = w_0$. 

149
The components of the first moment are calculated from the equations (Manning, 1968)

\[ \langle x \rangle = N_{\text{jump}} a (p_{x+} - p_{x-}) , \]  
(4.54a)

\[ \langle y \rangle = N_{\text{jump}} a (p_{y+} - p_{y-}) , \]  
(4.54b)

where the number of jumps \( N_{\text{jump}} \) is equal to the product \( N_{\text{jump}} = wt \) and \( (x+, y+, x-, y-) \) are the directions \((F_x \hat{x}, F_y \hat{y}, -F_x \hat{x}, -F_y \hat{y})\). Using eqs. (4.52)-(4.54) and the result of the first moment analysis (4.45) gives the components of the drift velocity as

\[ u_{d,x} = 2aC \exp\left( -\frac{Q^o}{kT} \right) \sinh\left( \frac{F_x a}{2kT} \right) , \]  
(4.55a)

\[ u_{d,y} = 2aC \exp\left( -\frac{Q^o}{kT} \right) \sinh\left( \frac{F_y a}{2kT} \right) . \]  
(4.55b)

The component \( \langle x^2 \rangle \) of the second moment tensor is calculated from the equation (Manning, 1968)

\[ \langle x^2 \rangle = N_x a^2 + \frac{N_x - 1}{N_x} \langle x \rangle^2 , \]  
(4.56)

where \( N_x \) is the number of jumps in the \( x- \)direction. In the limit \( t \to \infty \), \( \frac{N_x - 1}{N_x} \approx 1 \) and eq. (4.56) gives

\[ \langle x^2 \rangle - \langle x \rangle^2 = N_x a^2 , \]  
(4.57)

and \( N_x = w_x t \) and \( w_x = w_{x+} + w_{x-} \). The analysis is exactly the same in the \( y- \)direction.

The crystallographic axes of the square and cubic lattices are the principal axes of the second central moment tensor; as a result, the diffusivity tensor is diagonal. The symmetry properties of the lattice, or the periodic medium in general, are responsible for the structure of the material property tensors (Manning, 1968; Brenner, 1980). Equations (4.47), (4.56), (4.57) and the result of the second moment analysis (4.46b) give the components of the diffusivity tensor as

\[ D_{xx} = a^2 C \exp\left( -\frac{Q^o}{kT} \right) \cosh\left( \frac{F_x a}{2kT} \right) , \]  
(4.58a)

\[ D_{yy} = a^2 C \exp\left( -\frac{Q^o}{kT} \right) \cosh\left( \frac{F_y a}{2kT} \right) . \]  
(4.58b)
In the case of a zero applied force, eq. (4.58) gives the well known result of random walk diffusion
\[ D_{xx} = D_{yy} = a^2 w_0 \exp\left(-\frac{Q^o}{kT}\right). \] (4.59)

A very interesting result of the analytical solution for interstitial diffusion with constant force comes from the comparison of eqs. (4.55) and (4.58) to form the ratio
\[ \frac{D_{ij}}{u_{d,j}} = \frac{a \cosh\left(\frac{F_j a}{2kT}\right)}{2 \sinh\left(\frac{F_j a}{2kT}\right)}; \quad j = x, y, \] (4.60)

which proves that Einstein's relation (Einstein, 1905) \( u_d = D \cdot F/kT \) is not generally valid. Einstein's equation is valid in the asymptotic limit of low forces or high temperatures, where the arguments of the hyperbolic functions are much less than unity.

4.1.4.2. Computational Results

The Monte Carlo (MC) method introduced in Section 4.1.2 is used to calculate the drift velocity and the diffusivity tensor of interstitials in a two-dimensional square lattice under the action of a constant force \( F \). Interstitial random walk diffusion in a two-dimensional square lattice and a simple cubic lattice is used as a test for the Monte Carlo computations. An energy barrier of \( Q^o = 10 \text{ kJ/mol} \), typical of the activation energy for light interstitials in metals (Kirchheim, 1987), has been used over the range of temperatures from 300 K to 1000 K.

The variance of the calculations for the first and second moments increases with the number of MC steps. The moments are calculated by averaging over \( M \) particle trajectories. The time dependence of the moments is initially examined using a long time run with a duration of \( 10^5 \) MC steps. The components of the first moment fluctuate about the average zero value. The components of the second moment fluctuate about a straight line, that is established after very short times. Averaging over different starting times (Jacoboni and Reggiani, 1983; Kehr and Binder, 1984) is used to improve the accuracy of the calculations. Each long trajectory is divided into \( m \) subtrajectories. The duration of the \( k \)th subtrajectory is \( t_k = t_\ast / m \), where \( t_\ast \) is the total simulation time. The initial position of
the interstitial in the $k+1$ subtrajectory is its final position in the $k$th subtrajectory. The result for the diffusivity tensor is computed by time averaging over the ensemble averages constructed over each subtrajectory as

$$
\bar{D} = \frac{1}{m} \sum_{k=1}^{n} D_k, \quad (4.61a)
$$

where $D_k$ is calculated using the simple finite difference scheme

$$
D_k = \frac{1}{2t_k} \langle rr \rangle_k. \quad (4.61b)
$$

The accuracy of the simulation is proportional to $1/\sqrt{Mm}$. A number of $M = 1000$ particle trajectories and $m = 100$ subtrajectories with a duration of $t_k = 1000$ MC steps is used for all calculations.

The variation of the average value of the components of $D$ with respect to the number of subtrajectories $m$ is presented in Fig. 4.3 for both the square and the simple cubic lattices at temperature $T = 1000$ K. For the square lattice, the off-diagonal element $D_{xy}$ is zero, as expected. The diagonal elements converge to the same final average value, within the limits of statistical uncertainty, thereby proving that the diffusivity tensor is isotropic and verifying the low bias of the random number generator. In addition, comparison of the numerical results for the two-dimensional and the three-dimensional random walk problem verifies the correct introduction of the time variable. Using the numerical results shown in Fig. 4.3 gives

$$
\frac{D_{2D}}{D_{3D}} = 1.50 \pm 0.01, \quad (4.62a)
$$

when the same activation energy and lattice parameter are used for both problems. The analytical result, eq. (4.59), gives

$$
\frac{D_{2D}}{D_{3D}} = \frac{w_{0,2D}}{w_{0,3D}} = \frac{Z_{3D}}{Z_{2D}} = \frac{6}{4}, \quad (4.62b)
$$

according to eq. (4.21).
Figure 4.3. Convergence of the components $D_{ij}$ to a final equilibrium value as a function of time. The time evolution is given by the number of subtrajectories, each equal to 1000 MC steps. The temperature for the simulation is $T = 1000$ K. Results are shown for (a) a two-dimensional square lattice and (b) a simple cubic lattice.
The correct value of $Q^o = 10 \text{ kJ/mol}$ in the diffusivity expression is recovered from the results of Fig. 4.3 using $a = 1$, where the lattice parameter is the unit length for the simulation. The value of $Q^o = 10 \text{ kJ/mol}$ in eq. (4.59) also has been checked to be equal to the input value from the Arrhenius plots for the temperature dependence of two-dimensional and three-dimensional diffusion coefficients.

Interstitial diffusion with constant force $F$ is simulated in a two-dimensional square lattice, where the applied force is represented as $F = F_x\hat{x} + F_y\hat{y}$. The value of $Q^o$ and the temperature range are set equal to the values used in the random walk problems. The barriers $Q_{ij}$ are calculated from eq. (4.19). The components of the force are varied in the ranges $-1 \leq F_x \leq 0$ and $-1 \leq F_y \leq +1$ in units of $Q^o/a$. The drift velocity and diffusivity components are calculated by the equations

$$\tilde{u}_d = \frac{1}{m} \sum_{k=1}^{m} u_{d,k}, \quad \tilde{D} = \frac{1}{m} \sum_{k=1}^{m} D_k,$$

(4.63a)

where

$$u_{d,k} = \langle r \rangle_k, \quad D_k = \frac{1}{2t_k} \left( \langle r^2 \rangle - \langle r \rangle \langle r \rangle \right)_k.$$

(4.63b)

The evolution of the first moments as a function of time is shown in Figs. 4.4a and 4.4b, respectively, for $T = 1000$ K and $T = 300$ K and for $F_x = -0.5 \frac{Q^o}{a}$ and $F_y = -0.2 \frac{Q^o}{a}$. The first moments are linear in time with almost no fluctuations, because of the relatively large applied force. The second central moment components are given in Figs. 4.5a and 4.5b for the same cases. The off-diagonal components of the second central moment tensor are computed to be zero within the limits of statistical uncertainty.

The most interesting result is the differing behavior of the anisotropy as a function of temperature. At high temperature ($T = 1000$ K) the anisotropy introduced by $F$ is small and the diffusivity is almost isotropic. At $T = 300$ K the anisotropy introduced by $F$ is significant, as measured by the ratio $|D_{xx} - D_{yy}|/D_{xx}$ calculated from the slopes of the two fluctuating straight lines in Fig. 4.5. The anisotropy introduced at low temperatures also is clearly seen in Fig. 4.6, where the variation of the components $\{D_{ij}\}$ with the number of subtrajectories $m$ is displayed.
Figure 4.4. Time dependence of the components of the first moment for diffusion with constant force in a square lattice for $F_x = -0.5 \, Q^o / a$ and $F_y = -0.2 \, Q^o / a$: (a) $T = 1000$ K and (b) $T = 300$ K.
Figure 4.5. Time dependence of the components of the second central moment for diffusion with constant force in a square lattice for $F_x = -0.5 \, Q^o/a$ and $F_y = -0.2 \, Q^o/a$: (a) $T = 1000 \, \text{K}$ and (b) $T = 300 \, \text{K}$. 
Figure 4.6. Convergence of the components $D_{ij}$ to equilibrium values as a function of time for $F_x = -0.5 \, Q^o/a$ and $F_y = -0.2 \, Q^o/a$: (a) $T = 1000$ K and (b) $T = 300$ K.
The temperature dependence of the diffusivity components is presented in the Arrhenius plots in Fig. 4.7 for ranges of the force components \( F_x \) and \( F_y \). The results of the MC simulation are in complete agreement with the analytical results (4.58). The Arrhenius behavior is obvious. Finally, the dependence of the diffusivity components on the magnitude of the force \( \mathbf{F} \) is presented in Fig. 4.8 for \( T = 300 \text{ K} \). The diffusivity components are plotted as functions of \( F_x \) for constant \( F_y \). Anisotropy becomes more important as the magnitudes of \( (F_x, F_y) \) increase. For \( F_x = F_y \), the diffusivity is isotropic. Again, the agreement between the Monte Carlo results and the analytical solution is very good. For the comparison between analytical and computational results the lattice parameter \( a \) is used as the unit length \( (a = 1) \) and the value \( w_0 = 1/4 \) is used as the definition of the time unit for the simulation.

4.1.5. Extension to Realistic Systems and Inhomogeneous Forces

The results of Section 4.1.4 for interstitial impurity migration under constant force in a two-dimensional square lattice are easily extended to other cubic lattice structures. For example, the extension of equations (4.55) and (4.58) to a three-dimensional simple cubic lattice is straightforward. For other cubic structures the extensions require the incorporation of the geometrical details of the lattice for the derivation of constitutive equations that relate drift velocities and diffusivities to applied or internal forces and temperature. A detailed analysis for the diamond-cubic structure is given in Section 4.2 where the migration of interstitial oxygen is analyzed under the force field created by glide dislocations.

The extension for inhomogeneous force fields depends on the strength and the range of the interaction with the interstitial impurity in the lattice. The extension of the results of Section 4.1.4 to inhomogeneous forces is rigorous if higher order drift terms in eq. (4.14) are significantly smaller than the first-order terms, which are the only nonzero drift terms in the case of a constant force. In such a case, eqs. (4.55) and (4.58), or the analogous equations in other cubic structures, give local values for the drift velocity and the diffusivity, respectively, at the appropriate values of the force \( \mathbf{F} = \mathbf{F}(r) \).
Figure 4.7. Arrhenius plots for the temperature dependence of the diagonal components of the diffusivity tensor: (a) $F_x = -0.5 \, Q^0/a$ and $F_y = -0.2 \, Q^0/a$; (b) $F_x = -0.5 \, Q^0/a$ and $F_y = 0$; (c) $F_x = -0.5 \, Q^0/a$ and $F_y = 0.5 \, Q^0/a$. The results of the Monte Carlo simulation for $D_{xx}$ ($\bigcirc$) and $D_{yy}$ (+) are compared with the results of eqs. (49a) (---) and (49b) (- - - -). The solid line (-----) in each plot gives the two-dimensional random walk result (50).
Figure 4.8. Dependence of the diffusivity components on the magnitude of the component $F_x$. The results of the Monte Carlo simulation for $D_{xx}$ (○) and $D_{yy}$ (△) are compared with the results of eqs. (49a) (- - -) and (49b) (---). Calculations are shown for $F_y$ equal to 0.1, 0.3, 0.5 and 0.8 $Q^0/a$.
The validity of the extension to inhomogeneous force fields is demonstrated for the migration of an electrically inactive interstitial impurity around an edge dislocation in a face-centered cubic lattice. The first order drift term \( \nabla U \cdot (\mathbf{r}_s - \mathbf{r}_e) \) is compared with the second order drift term \( \frac{1}{2} \nabla \nabla U : (\mathbf{r}_s - \mathbf{r}_e)(\mathbf{r}_s - \mathbf{r}_e) \). The interaction energy \( U \) in this case is

\[
U = A_d b \frac{\sin \theta}{r},
\]

(4.64)

where \( b \) is the magnitude of the Burgers vector of the dislocation. For a face-centered cubic lattice and an undissociated dislocation \( b = a \sqrt{2}/2 \), where \( a \) is the lattice parameter. The distance between two nearest neighbor sites is equal to \( b \) and yields

\[
\mathbf{r}_s - \mathbf{r}_e = \frac{1}{2} b \hat{\mathbf{e}}_x.
\]

(4.65)

Therefore, the terms to be compared are \(-\frac{1}{2} F_x b^2 \) and \(-\frac{1}{8} \frac{\partial F_x}{\partial x} b^2 \) for jumps in the \( x \)− direction and \(-\frac{1}{2} F_y b^2 \) and \(-\frac{1}{8} \frac{\partial F_y}{\partial y} b^2 \) for jumps in the \( y \)− direction. In the \( (001) \) plane the cartesian axes \( x \) and \( y \) coincide with the crystallographic axes \([\overline{1}0] \) and \([10] \), respectively.

Using \( \mathbf{F} = -\nabla U \) and eq. (4.64) gives

\[
F_x = \frac{A_d b}{r^2} \sin 2\theta, \quad F_y = -\frac{A_d b}{r^2} \cos 2\theta,
\]

(4.66a)

\[
\frac{\partial F_x}{\partial x} = -2 \frac{A_d b}{r^3} \sin 3\theta, \quad \frac{\partial F_y}{\partial y} = 2 \frac{A_d b}{r^3} \sin 3\theta.
\]

(4.66b)

The ratio of the absolute values of terms at distances \( r = na \), where \( n \) is an integer are

\[
\left| \frac{\frac{1}{8} \left( \frac{\partial F_x}{\partial x} \right) b^2}{\frac{1}{2} F_x b} \right| = \frac{1}{2} \left( \frac{b}{r} \right) \left| \frac{\sin 3\theta}{\sin 2\theta} \right| = \frac{\sqrt{2}}{4n} \left| \frac{\sin 3\theta}{\sin 2\theta} \right|,
\]

(4.67a)

for jumps in the \( x \)− direction, and

\[
\left| \frac{\frac{1}{8} \left( \frac{\partial F_y}{\partial y} \right) b^2}{\frac{1}{2} F_y b} \right| = \frac{1}{2} \left( \frac{b}{r} \right) \left| \frac{\sin 3\theta}{\cos 2\theta} \right| = \frac{\sqrt{2}}{4n} \left| \frac{\sin 3\theta}{\cos 2\theta} \right|,
\]

(4.67b)

for jumps in the \( y \)− direction. Therefore, setting the trigonometric functions of the right hand sides of eqs. (4.67) equal to unity, the ratios are equal to 0.18 for \( n = 2 \) and 0.07
for \( n = 5 \) falling in general as \( 1/r \). The first order drift terms obviously dominate and the truncation of the Taylor series is justified.

4.2. Migration of Interstitial Oxygen in the Dislocation Field of a Silicon Lattice

Quantitative analysis of oxygen gettering by dislocations is based on the solution of the transport equation (4.29) for interstitial oxygen with the appropriate boundary conditions (Sumino, 1987a). Previous numerical solutions of this problem put the dislocation at the center of an imaginary square lattice and use an isotropic diffusivity and a drift velocity that satisfies Einstein’s equation (Einstein, 1905). In this section a constitutive model for the drift velocity and diffusivity of oxygen around 60° glide dislocations is derived following the analysis of Section 4.1.1.

4.2.1. Analysis of Oxygen Diffusion Around a 60° Glide Dislocation in Silicon

The dislocation field is modelled as a regular array of parallel 60° dislocations with dislocation density \( N \). If the dislocations lie along the \( z \)-axis of a cartesian coordinate system, oxygen migration driven by the force field generated by the dislocation occurs in the \( xy \) plane. The analysis is similar for a regular array of screw dislocations.

The geometry of the interstitial diffusion process is determined by the glide dislocation configuration in the diamond crystal structure (Sumino and Harada, 1981; Alexander, 1986). A schematic representation of the glide dislocation with respect to the conventional cubic unit cell of the lattice is shown in Fig. 4.9a. In the figure, the line OE is the 60° dislocation which lies on the (\( \bar{1}11 \)) plane (OAE) and glides along the [\( \bar{1}0\bar{1} \)] direction (AO). The Burgers vector of the dislocation is \( b = \frac{a}{2}[\bar{1}0\bar{1}] \). The dislocation is perpendicular to the (110) plane (ABCD), which is the plane where diffusion occurs. The cartesian coordinate system \( xy \) in the (110) plane is determined by the dislocation velocity \( \mathbf{v} \), which is parallel to the [\( \bar{1}0\bar{1} \)] direction. The coordinate axis \( x \) is the projection of the [\( \bar{1}0\bar{1} \)] direction in the (110) plane, i.e. line AL in Fig. 4.9a. Therefore, the \((x, y)\) cartesian coordinates are \( x \equiv [\bar{1}1\bar{2}] \) and \( y \equiv [\bar{1}11] \), as shown in Fig. 4.9b.
Figure 4.9. (a) Schematic representation of a $60^\circ$ glide dislocation (line OE) with respect to the conventional cubic unit cell of the diamond lattice. (b) Two-dimensional hexagonal cells of silicon atoms in the (110) plane (ABCD) that is perpendicular to the dislocation line.
The silicon atoms in the (110) plane form a two-dimensional lattice that consists of hexagonal cells (see the hexagon MPCSRQ in Fig. 4.9b) with edges of length \( MQ = CS = \frac{a}{4} \) and \( MP = PC = QR = RS = \frac{a\sqrt{3}}{4} \). The lattice parameter \( a \) is equal to \( 5.43 \times 10^{-10} \) m for silicon. The length MP is equal to the length of a bond in the tetrahedral (covalent) bonding of the diamond structure. The angles \( \theta \) and \( \phi \) of the hexagon are \( \theta = 109.47^\circ \) and \( \phi = 125.26^\circ \).

The interstitial atom is put at the center of the hexagonal cell, which is represented in Fig. 4.10a. The interstitial atom at site O has four nearest neighbors at distances \( l = OA = OB = OC = OD = \frac{a}{2}\sqrt{3/2} \). Migration occurs by jumps from the site O to one of the four nearest neighbors. The distance \( l' = OE = OG = a\sqrt{2}/2 > l \) and the barrier for the corresponding jumps is higher than the barrier for jumps to the four nearest neighbors in the absence of any forces.

The analysis presented in Section 4.1.4 for interstitial diffusion in a two-dimensional lattice is applied with the crystallographic axes of the interstitial lattice taken as \( x_1 \equiv [\overline{1}12] \) and \( x_2 \equiv [\overline{1}12] \) as shown in Fig. 4.10b. The corresponding unit vectors are

\[
\begin{align*}
\hat{x}_1 &= \hat{x}, \\
\hat{x}_2 &= -\frac{1}{3}\hat{x} + \frac{2\sqrt{2}}{3}\hat{y}.
\end{align*}
\]  

(4.68a) \hspace{1cm} (4.68b)

### 4.2.2. Constitutive Model for the Drift Velocity and the Diffusivity

The interaction force \( \mathbf{F}(r) = -\nabla U(r) \) is calculated from the interaction potential between a point defect and a 60° dislocation. Taking the interaction to be elastic (Hirth and Lothe, 1982) gives

\[
U(r) = \frac{\sqrt{3}}{2} A_d b \sin \theta \frac{\sin \theta}{r},
\]

(4.69)

where the point defect is assumed to be spherical, \( A_d \approx 0.5 \) eV for oxygen in silicon (Sumino and Imai, 1983) and \( r \) is the distance between the point defect and the center of the dislocation core. Transforming the elastic interaction potential (4.69) from polar to
Figure 4.10. (a) The interstitial lattice structure of the two-dimensional silicon lattice in the (110) plane. (b) The crystallographic axes of the two-dimensional interstitial lattice.
cartesian coordinates gives the components of the force $F = F_x \hat{x} + F_y \hat{y}$ as

$$F_x = \frac{\sqrt{3}}{2} A_d b \frac{\sin 2\theta}{r^2}, \quad (4.70a)$$

$$F_y = -\frac{\sqrt{3}}{2} A_d b \frac{\cos 2\theta}{r^2}, \quad (4.70b)$$
or in the $(x_1, x_2)$ directions, where $F = F_1 \hat{x}_1 + F_2 \hat{x}_2$, as

$$F_1 = F_x + \frac{\sqrt{2}}{4} F_y, \quad (4.71a)$$

$$F_2 = \frac{3\sqrt{2}}{4} F_y. \quad (4.71b)$$

Application of the results of Section 4.1.4 is now straightforward. Dimensionless force components and activation energies scaled with the thermal energy are defined as

$$f_1 \equiv \frac{F_1 l}{2kT}, \quad f_2 \equiv \frac{F_2 l}{2kT}, \quad q \equiv \frac{Q^0}{kT}, \quad f_{max} \equiv \frac{\max\{F_1, F_2\} l}{2kT}. \quad (4.72)$$

The drift velocity and the diffusivity tensor are written in the transformed coordinate system $(x_1, x_2)$ as $u_d \equiv u_1 \hat{x}_1 + u_2 \hat{x}_2$ and $D \equiv D_{11} \hat{x}_1 \hat{x}_1 + D_{12} \hat{x}_1 \hat{x}_2 + D_{21} \hat{x}_2 \hat{x}_1 + D_{22} \hat{x}_2 \hat{x}_2$. The off-diagonal terms ($D_{12}, D_{21}$) are equal to zero, because the component $(\langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle)$ of the second central moment tensor is zero from the symmetry of the lattice; therefore, $D \equiv D_{11} \hat{x}_1 \hat{x}_1 + D_{22} \hat{x}_2 \hat{x}_2$. The directions $(\hat{x}_1, \hat{x}_2)$ are not orthogonal and hence are not principal axes of $D$. According to eqs. (4.55) and (4.58) the components of the drift velocity and diffusivity are

$$u_1 = 4lw_0 \exp(-q) \exp(f_{max}) \frac{\sinh(f_1)}{\cosh(f_1) + \cosh(f_2)}, \quad (4.73a)$$

$$u_2 = 4lw_0 \exp(-q) \exp(f_{max}) \frac{\sinh(f_1)}{\cosh(f_1) + \cosh(f_2)}, \quad (4.73b)$$

$$D_{11} = 2l^2 w_0 \exp(-q) \exp(f_{max}) \frac{\cosh(f_1)}{\cosh(f_1) + \cosh(f_2)}, \quad (4.73c)$$
\[ D_{22} = 2l^2 w_0 \exp(-q) \exp(f_{\text{max}}) \frac{\cosh(f_2)}{\cosh(f_1) + \cosh(f_2)}. \] (4.73d)

Transforming these results to the cartesian coordinate system \((x, y)\), gives the components of the drift velocity \(u_d = u_{d,x} \hat{x} + u_{d,y} \hat{y}\) as

\[ u_{d,x} = u_1 - \frac{1}{3} u_2, \] (4.74a)

\[ u_{d,y} = \frac{2\sqrt{2}}{3} u_2, \] (4.74b)

and the diffusivity tensor as

\[ D = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} = \begin{pmatrix} D_{11} + \frac{1}{9} D_{22} & -\frac{2\sqrt{2}}{9} D_{22} \\ -\frac{2\sqrt{2}}{9} D_{22} & \frac{8}{9} D_{22} \end{pmatrix} \] (4.74c)

This tensor is symmetric and positive definite with determinant \(\text{det}(D) = \frac{8}{9} D_{11} D_{22}\).

### 4.2.3. Diffusivity Data for Oxygen Diffusion in the (110) Plane

The analysis is completed by providing diffusivity data for interstitial oxygen diffusion in the (110) plane of the silicon lattice. The results are based on the constitutive model, eqs. (4.73) and (4.74), and isotropic diffusivity data for diffusion in a three-dimensional lattice.

In the absence of an external force \(\mathbf{F}\), the diffusivity of oxygen in silicon is given by the diffusion coefficient (Mikkelsen, 1982; Stavola et al., 1983; Inoue et al., 1987)

\[ D = 0.091 \exp\left(-\frac{2.4\text{eV}}{kT}\right) \text{cm}^2/\text{s}, \] (4.75)

in a temperature range from 543 K to 1473 K. The attempt frequency \(w_0\) is computed by fitting this data to the result (4.59) using a value of \(a\) equal to the bond length for the three-dimensional silicon lattice \(l_3 = a\sqrt{3}/4 = 2.35 \times 10^{-10} \text{m}\); the result of this calculation is \(w_0 = 1.65 \times 10^{14} \text{s}^{-1}\).

The experimental results are transformed to the two-dimensional lattice assuming the same activation energy barrier \(Q^a = 2.4 \text{eV}\) and the same attempt frequency \(w_0\); these
assumptions are reasonable because the coordination number is the same for both lattices. Using the bond length \( l = \frac{a}{2} \sqrt{\frac{3}{2}} = 3.32 \times 10^{-10} \text{m} \) in eqs. (4.73) and (4.74) gives the values of the drift velocity and the diffusivity. The Arrhenius plot of the diffusivity components is shown in Fig. 4.11 for the values at the position \((r = 5a, \theta = \frac{\pi}{2})\) from the center of the dislocation core.

The components of the diffusivity tensor reach their maximum values next to the dislocation core, where the interaction forces (4.70) are maximum. These maximum values vary from 10 times to twice the isotropic value (4.75) as the temperature varies from 500 K to 1500 K. The ratio \( \frac{D_{zz}}{D_{yy}} \) is used as a measure of the anisotropy of the diffusivity tensor (61c). The maximum value of this ratio varies from 5.1 to 1.6 as the temperature varies from 500 K to 1500 K. The calculation of the drift velocity and the diffusivity of oxygen in the stress field created by 60° dislocations opens the way for precise analysis of oxygen transport around dislocations in silicon.

4.3. Oxygen Concentration Field Around Dislocations in Silicon

The effects of oxygen on dislocation mobility are the result of elastic interaction between dislocations and the impurity and the electrical interactions caused by the thermal donor characteristics of SiO\(_x\) complexes. Also, oxygen forms aggregates that decorate the dislocation and are thought to inhibit the onset of dislocation motion (Sumino, 1987a). In general, the relative importance of these modes for interaction of oxygen with dislocations depends on the thermal and strain histories of the sample. Because the interstitial oxygen level introduced during Czochralski crystal growth (1 \( \times 10^{18} \) atoms/cm\(^3\)) is above the saturation limit for temperatures below 1500 K (Inoue et al., 1987), the oxygen precipitates out of solid solution as the material is processed at lower temperatures. Oxygen clusters are formed, whether or not dislocations are present in the sample. If they are, the clusters have a high affinity for forming along the dislocation (Sumino, 1987a). In fact, infrared absorption measurements indicate that all of the oxygen in a typical specimen may be gettered to the dislocations over time when the dislocation density is high (Yonenaga and Sumino, 1985). Both the dispersed and aggregated oxygen in the dislocation core affect
Figure 4.11. Arrhenius plot for the temperature dependence of the diffusivity of interstitial oxygen in silicon in terms of the values of the components of the diffusivity tensor for diffusion in the (110) plane. Results are determined at the position \( r = 5a, \theta = \pi/2 \) from the center of the dislocation core, as derived from eqs. (60) and (61c).
the ability of the dislocation to glide under an applied stress. The analysis of the effects of oxygen on dislocation motion is presented in detail in Chapter 5.

The analysis of this section concentrates on the concentration field of dissolved oxygen outside of the cores of immobile or gliding dislocations. The analysis of the anisotropic migration process is based on the constitutive equations derived in Section 4.2. Cottrell (1953) first realized that the elastic stress field caused by the interaction between a dislocation and an impurity atom leads to migration of the impurity to the dislocation. This driving force, coupled with the diffusion of the impurity into the core causes enhanced oxygen concentrations or atmospheres near the dislocation and a drag force on the dislocation motion. The analysis of oxygen migration to dislocations extends the pioneering work of Cottrell and Bilby (1949) and Ham (1959) of impurity migration around immobile dislocations to dislocations moving with fixed velocity. Cottrell and Jawson (1949) also considered this problem, but limited their study to low dislocation velocities, where a perturbation analysis constructed about the solution for the immobile defect is valid. The only previous study of impurity atmospheres near mobile dislocations are the limited results of Yoshinaga and Morozumi (1971) for carbon in vanadium.

The mathematical formulation of the model for the oxygen concentration field is presented in Section 4.3.1. The numerical method for the solution of the transport problem for a mobile dislocation is described in Section 4.3.2. Sample oxygen distributions around 60° dislocations are presented in Section 4.3.3 to demonstrate the formation of oxygen atmospheres around immobile and gliding dislocations and the evolution of the atmospheres as the dislocation velocity is increased.

4.3.1. Mathematical Formulation

Three assumptions are introduced in the analysis of oxygen migration. First, it is assumed that no oxygen precipitation occurs outside of the dislocation core. A typical oxygen concentration for Czochralski grown silicon is \( c = 10^{18} \text{ cm}^{-3} \) (Sumino, 1987a; Brown, 1988). The solubility of oxygen in silicon depends on temperature according to
(Inoue et al., 1987)

\[ c_{eq} = 2.00 \times 10^{21} \exp\left(-\frac{1.032 \text{ eV}}{kT}\right) \text{ cm}^{-3}. \]  

(4.76)

Therefore, the condition \( c < c_{eq} \) is valid in the temperature range \( 1500 \text{ K} < T < T_m \), where \( T_m = 1683 \text{ K} \) is the melting point of silicon. The analysis is extrapolated to temperatures lower than 1500 K for regions of the crystal outside dislocation cores. The formation of oxygen precipitates outside the core is neglected because silicon samples used for dislocation mobility and dynamic deformation experiments are preannealed at \( T = 1573 \text{ K} \) to dissolve oxygen precipitates that have formed during the cooling of the crystal after growth (Yonenaga et al., 1984; Sumino, 1987a). As soon as dislocations are introduced in the sample, the interaction between oxygen atoms and dislocations dominates over interactions between the interstitial oxygen atoms in the lattice and precipitate formation is avoided in favor of oxygen gettering to the dislocation core. Even if a very small number of precipitates exists outside the core, these migrate by a two-step process at a much slower rate than interstitial oxygen (Needels et al., 1991) and may be neglected compared to the transport of dilute distributed interstitial oxygen.

Second, it is assumed that the interaction potential between interstitial oxygen atoms and the 60° dislocation is given by the first-order elastic potential due to the size misfit of oxygen in the silicon matrix (Bullough and Newman, 1970; Hirth and Lothe, 1982; Nabarro, 1987). This assumption is satisfactory for two reasons. First, the first-order interaction is much larger than higher-order elastic interactions for dislocations with edge components (Bullough and Newman, 1970). Also, dissolved oxygen is an electrically inactive impurity in silicon (Bullough and Newman, 1970; Mikkelsen, 1986). The interaction potential between an oxygen atom and the 60° dislocation is given from dislocation theory (Hirth and Lothe, 1982; Nabarro, 1987) by equation (4.69)

\[ U(r) = \frac{\sqrt{3}}{2} A_d b \frac{\sin \theta}{r}, \]  

(4.69)

where \((r, \theta)\) are cylindrical polar coordinates, and \(b\) is the magnitude of the Burgers vector of the dislocation, and the dislocation lies along the \(z\)-axis of the cylindrical coordinate system. The constant \(A_d\) is calculated from dislocation theory using the elastic constants of

171
silicon and the sizes of the oxygen atom and the matrix atom (Bullough and Newman, 1970; Yoshinaga and Morozumi, 1971). The form of the interaction potential (4.69) implies that oxygen migration occurs in two dimensions for an initially uniform oxygen level because the force component in the z-direction is equal to zero. Therefore, there is no initial driving force for oxygen transport in the z-direction when the oxygen field is uniform in this direction.

Also, the potential (4.69) does not attempt to capture the core region of the dislocation and therefore is valid only in a region $r \geq r_0$ away from the center of the dislocation core. Recent atomistic calculations for the structure of the dislocation core (Nandedkar and Narayan, 1987) have determined the core radius $r_0$ and verified the validity of elasticity theory away from the core. For a $60^\circ$ dislocation in silicon the core radius was calculated to be approximately $0.9a \simeq 5 \times 10^{-10}$ m, where $a$ is the lattice parameter ($5.43 \times 10^{-10}$ m) for silicon. This leads to the third assumption that the size of the dislocation core may be specified.

4.3.1.1. Transport Equation and Boundary Conditions

The conservation equation for the concentration $c(r,t)$ of an impurity based on equation (4.29) is written in the frame of reference of a moving dislocation as

$$\frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{J} \equiv \nabla \cdot [(\mathbf{v} - \mathbf{u}_d)c + \mathbf{D} \cdot \nabla c], \tag{4.77}$$

where $\mathbf{J}$ is the total impurity flux, $\mathbf{v}$ is the dislocation velocity, $\mathbf{u}_d$ is the drift velocity of the impurity due to the force exerted by the dislocation and $\mathbf{D}$ is the diffusivity tensor of the impurity in the stress field created by the dislocation.

Yoshinaga and Morozumi (1971) have shown that the migration of impurities around dislocations reaches a steady state, irrespective of whether the initial concentration profile is uniform or at equilibrium with a stationary defect. This steady state is reached within times of $O(10^{-2}$ s), where the precise value depends on the dislocation velocity. The steady state concentration of the impurity concentration is described by

$$\nabla \cdot [(\mathbf{v} - \mathbf{u}_d)c + \mathbf{D} \cdot \nabla c] = 0, \tag{4.78}$$

172
and is the basis for our analysis. Equation (4.78) is written in dimensionless form by scaling the oxygen concentration with the uniform initial concentration $c_0$, lengths with $a$, velocity components with a characteristic drift velocity $v_0$, and components of the diffusivity tensor $D$ with the isotropic diffusivity $D$ of oxygen in silicon. Assuming that the dislocation moves in the $x-$direction of a cartesian coordinate system, the dimensionless form of eq. (4.78) is

$$\hat{\nabla} \cdot \left[ Pe_d (\mathbf{V} - \mathbf{u}_d) \hat{c} + \hat{\Delta} \cdot \hat{\nabla} \hat{c} \right] = 0,$$

where $\hat{\nabla}$ is the dimensionless gradient operator, $\mathbf{V} = \mathbf{v}/v_0$, $\mathbf{u}_d = \mathbf{u}_d/v_0$, $\hat{c} = c/c_0$, $\hat{\Delta} = D/D$, $Pe_d$ is the drift Peclet number that scales oxygen drift caused by the interaction with the dislocation with diffusion

$$Pe_d \equiv \frac{v_0 a}{D},$$

and $\hat{x}$ is the unit vector in the direction of the glide motion.

The solution of the linear, elliptic partial differential equation (4.79) requires the definition of a domain and appropriate boundary conditions. The domain is defined in terms of the dislocation density $N$ of the matrix by placing the dislocation at the center of a square with size $L = 1/\sqrt{N}$, as shown in Fig. 4.12a.

Typical dislocation densities in semiconductor materials are in the range $10^3 \leq N \leq 10^5$ cm$^{-2}$, either as grown from the melt or during the later stages of thermal processing (Sumino, 1987a; Jordan and Parsey, 1988). Therefore, the size of the periodic structure is $3.2 \times 10^{-5}$ m $\leq L \leq 3.2 \times 10^{-4}$ m, so that $L$ is large compared to the dislocation core $r_0 = O(a)$; typically $L/r_0 \simeq 10^5 - 10^6$. This large difference suggests the introduction of a reduced computational domain that simplifies the solution of eq. (4.79). The reduced domain is taken to be the cylindrical annular region $r_0 \leq r \leq R$ shown in Fig. 4.12b. We take the outer radius $R$ as the distance from the dislocation core at which the impurity-dislocation interaction is equal to one percent of the interaction at the core radius $r = r_0$, i.e. $R = 100$ $r_0$. For a $60^\circ$ dislocation it is reasonable to assume $r_0 \simeq a$ and the annular domain becomes $a \leq r \leq 100$ $a$ or $1 \leq \hat{r} \leq \hat{R}$, with $\hat{R} = 100$.

The boundary condition on the oxygen concentration at $r = r_0$ models the incorporation or rejection of oxygen from the dislocation core. The choice of boundary condition
Figure 4.12. Domains for modelling oxygen field about (a) a regular array of dislocations and (b) an isolated dislocation.
has been discussed extensively by Bullough and Newman (1970). In many circumstances
the core can be assumed to be a perfect sink for the impurity atoms, so that the boundary
condition \( c(r_0, \theta) = 0 \) is appropriate (Bullough and Newman, 1970). This condition has
been discussed at length by Balluffi and Granato (1979) and is appropriate when the
impurity concentration near the core is small. At long times it is reasonable to assume
that the oxygen concentration in the core will reach local equilibrium, so that the net flux
is zero

\[
J \cdot n|_{r=r_0} = 0, \tag{4.81}
\]

where \( n \) is the unit normal to the circle \( (r = r_0) \). This assumption of local equilibrium was
used by Yoshinaga and Morozumi (1971). Boundary condition (4.81) is further justified by
the formation of oxygen precipitates inside the dislocation core and expresses the condition
of oxygen saturation inside the core and the state of dynamic equilibrium between dissolved
interstitial oxygen outside the core and oxygen precipitates inside the core.

The concentration outside the annular region \( r > R \) is taken to be the uniform value \( c_0 \),
so that the boundary condition at \( r = R \) is

\[
c(R, \theta) = c_0. \tag{4.82}
\]

Equation (4.79) and the boundary conditions (4.81) and (4.82) form the elliptic boundary-
value problem that is solved for the oxygen distribution for a given dislocation velocity
and applied force.

4.3.1.2. Constitutive Equations for Isotropic and Anisotropic Migration

The analysis of oxygen transport is presented for two constitutive models for the
diffusivity tensor \( D \) and the drift velocity \( u_d \) of oxygen.

The first model was introduced by Yoshinaga and Morozumi (1971) and uses a fictitious
square lattice with lattice constant \( l \) equal to the distance between the interstitial
and each of its four nearest neighbors. The diffusivity tensor is assumed to be isotropic
and independent of the local values of the interaction force; i.e.,

\[
D = DI \quad \text{and} \quad D = I, \tag{4.83}
\]

175
where the isotropic diffusivity has the experimental value \(4.75\). The drift velocity is related to the diffusivity and the local force field by Einstein's equation (Einstein, 1905) as

\[
u_d = \frac{DF}{kT}.
\]  \(4.84\)

Writing the drift velocity in terms of the force components \((F_x, F_y)\) given by eqs. \((4.70)\) reduces the transport equation \((4.79)\) to

\[
\hat{\nabla} \cdot \left[ P e_d (V \hat{x} - \frac{DF_x}{v_0 kT} \hat{x} - \frac{DF_y}{v_0 kT} \hat{y}) \hat{c} + \hat{\nabla} \hat{c} \right] = 0.
\]  \(4.85\)

The second constitutive model is the model for anisotropic migration developed in Section 4.2. The force is analyzed in the crystallographic directions \((x_1, x_2)\) \(\mathbf{F} \equiv F_1 \hat{x}_1 + F_2 \hat{x}_2\). The dimensionless forms for the drift velocity and the diffusivity \(\hat{u}_d\) and \(\Delta\) are taken from equations \((4.73)\) and \((4.74)\) as

\[
\hat{u}_d = \left[(\sinh f_1 - \frac{1}{3} \sinh f_2) \hat{x} + \frac{2\sqrt{2}}{3} \sinh f_2 \hat{y}\right] \frac{\exp(f_{\text{max}})}{\cosh f_1 + \cosh f_2},
\]  \(4.86a\)

\[
\Delta = \begin{pmatrix}
\cosh f_1 + \frac{1}{9} \cosh f_2 & -\frac{2\sqrt{2}}{9} \cosh f_2 \\
-\frac{2\sqrt{2}}{9} \cosh f_2 & \frac{8}{9} \cosh f_2
\end{pmatrix} \frac{\exp(f_{\text{max}})}{\cosh f_1 + \cosh f_2},
\]  \(4.86b\)

where the dimensionless force components \((f_1, f_2)\) are defined by eqs. \((4.71)\) and \((4.72)\) and \(f_{\text{max}} \equiv \max\{f_1, f_2\}\). The tensor \(\Delta\) is symmetric and positive definite.

The characteristic scales \(D\) and \(v_0\) used in the dimensionless form eq. \((4.79)\) are

\[
D = 2l^2 w_0 \exp\left(-\frac{Q^0}{kT}\right),
\]  \(4.87a\)

\[
v_0 = 4l w_0 \exp\left(-\frac{Q^0}{kT}\right),
\]  \(4.87b\)

where \(w_0 = 1.65 \times 10^{14} \text{ s}^{-1}\) and \(Q^0 = 2.4 \text{ eV}\) from Section 4.2.3. The drift Peclet number is \(Pe_d = 2a/l = 3.27\).
The transport equation (4.79) is solved in the annular domain using the cylindrical coordinate system shown in Fig. 4.12b. The dimensionless drift velocity \( \tilde{u}_d \) and diffusivity tensor \( \tilde{\Delta} \) are expressed in this system as

\[
\tilde{u}_d(r, \theta) = \tilde{u}_{d,r} \hat{r} + \tilde{u}_{d,\theta} \hat{\theta},
\]

(4.88a)

\[
\tilde{\Delta} = \Delta_{rr} \hat{r} \hat{r} + \Delta_{r\theta} \hat{r} \hat{\theta} + \Delta_{\theta r} \hat{\theta} \hat{r} + \Delta_{\theta \theta} \hat{\theta} \hat{\theta},
\]

(4.88b)

where \( \hat{r} \) and \( \hat{\theta} \) are the unit vectors in the radial and axial direction, respectively.

As first derived by Cottrell (1953), a closed form solution exists for the oxygen atmosphere in the limit of an immobile dislocation (\( v = 0 \) in eq. (4.78)) with an isotropic constant diffusivity \( D \). Then, the oxygen concentration is given by the Maxwellian atmosphere

\[
c = c_0 \exp\left( -\frac{U}{kT} \right),
\]

(4.89)

where \( U = U(r) \) is given by eq. (4.69). The derivation of (4.89) hinges on the validity of Einstein’s equation (4.84) for \( u_d \) in terms of a force \( F \) that is the gradient of a potential, i.e. \( F = -\nabla U \). At a constant temperature, the equipotential lines are circles that are symmetric about the \( y \)-axis of the cartesian coordinate system and the contours of equal concentration are circles about the \( y \)-axis.

Unfortunately, no closed-form solution exists either when the dislocation velocity is nonzero or when the expression for the anisotropic diffusivity is included. In the latter case, the drift velocity \( u_d \) cannot be written in terms of a gradient of a potential, as is easily seen by computing the curl of the effective force that corresponds to using an Einstein-like expression for the drift velocity with the anisotropic diffusivity expression (4.86).

4.3.2. Finite-Element Analysis

Because no closed form solution exists to eq. (4.79) when the dislocation moves with the velocity \( v = Vv_0 \hat{x} \), the concentration field is calculated numerically by the Galerkin finite element method. The application of Galerkin’s method to such a linear convection-diffusion problem is well documented in texts (Strang and Fix, 1973; Carey and Oden, 1983; Johnson, 1987); only the important details of our analysis are given here.
The annular region in Fig. 4.12 is discretized into a mesh of quadrilateral elements with $N_r$ and $N_\theta$ elements in the $r-$ and $\theta-$directions, respectively. The concentration field is approximated by an expansion using Lagrangian polynomial basis functions. Biquadratic polynomials $\{\Phi^j(\tilde{r}, \theta)\}$ defined on the elements are used to write the expansion for the concentration field as

$$\bar{c}(\tilde{r}, \theta) = \sum_{j=1}^{N_n} \alpha_j \Phi^j(\tilde{r}, \theta),$$

(4.90)

where $N_n$ is the number of nodes in the finite element mesh. Calculation of the coefficients $\{\alpha_j\}$ determines the concentration field $\bar{c}(\tilde{r}, \theta)$.

Galerkin’s method (Strang and Fix, 1973; Carey and Oden, 1983; Johnson, 1987) is used to reduce the weak form of the field equation and boundary conditions to a large set of sparse algebraic equations. Eq. (4.79) is weighted by the basis functions $\{\Phi^i(\tilde{r}, \theta)\}$ and it is integrated throughout the annular domain $\mathcal{D}$ to give

$$\int_{\mathcal{D}} \Phi^i \nabla \cdot \mathbf{J} dS = 0.$$  

(4.91)

Using the divergence theorem gives

$$\oint_{\partial \mathcal{D}} \mathbf{n} \cdot \Phi^i \mathbf{J} dl - \int_{\mathcal{D}} \mathbf{J} \cdot \nabla \Phi^i dS = 0,$$

(4.92)

where $\partial \mathcal{D}$ is the domain boundary and $\mathbf{n}$ is the unit vector normal to the boundary. The line integral of eq. (4.92) is decomposed as

$$\oint_{\partial \mathcal{D}} \mathbf{n} \cdot \Phi^i \mathbf{J} dl = \oint_{\partial \mathcal{D}_1} \Phi^i \mathbf{n} \cdot \mathbf{J} dl_1 + \oint_{\partial \mathcal{D}_2} \Phi^i \mathbf{n} \cdot \mathbf{J} dl_2,$$

(4.93)

where $\partial \mathcal{D}_1$ is the outer boundary, $\tilde{r} = 100$, and $\partial \mathcal{D}_2$ is the inner boundary, $\tilde{r} = 1$. Using boundary condition (4.81) and eqs. (4.92) and (4.93) gives

$$\oint_{\partial \mathcal{D}_1} \mathbf{n} \cdot \Phi^i \mathbf{J} dl_1 - \int_{\mathcal{D}} \mathbf{J} \cdot \nabla \Phi^i dS = 0.$$

(4.94)

The boundary condition at the outer boundary is the essential boundary condition $\bar{c}(100, \theta) = 1$. This is equivalent to setting the $\alpha_j = 1$ for all the nodes $j$ at the outer
boundary, due to the form of the Lagrangian basis functions. Therefore, the line integral of eq. (4.94) need not be calculated.

In cylindrical coordinates, the surface integral \( I_i = \int_D \mathbf{J} \cdot \nabla \Phi^i dS \) of eq. (4.94) is written as

\[
I_i = \int_1^{100} \int_0^{2\pi} \left[ \frac{\partial \Phi^i}{\partial \tilde{r}} \{ P e_d (V \cos \theta - \tilde{u}_{d,r}) \tilde{c} + \Delta_{rr} \frac{\partial \tilde{c}}{\partial \tilde{r}} + \frac{\Delta_{r\theta}}{\tilde{r}} \frac{\partial \tilde{c}}{\partial \tilde{\theta}} \} + \frac{1}{\tilde{r}} \frac{\partial \Phi^i}{\partial \tilde{\theta}} \{ P e_d (-V \sin \theta - \tilde{u}_{d,\theta}) \tilde{c} + \Delta_{r\theta} \frac{\partial \tilde{c}}{\partial \tilde{r}} + \frac{\Delta_{\theta\theta}}{\tilde{r}} \frac{\partial \tilde{c}}{\partial \tilde{\theta}} \} \} \right] \tilde{r} d\tilde{r} d\tilde{\theta}. \tag{4.95}\]

Using the Lagrangian approximation for the concentration field, eq. (4.90), and eq. (4.95) gives

\[
I_i = \sum_{j=1}^{N_n} \left[ \int_1^{100} \int_0^{2\pi} \left( \frac{\partial \Phi^i}{\partial \tilde{r}} \{ P e_d (V \cos \theta - \tilde{u}_{d,r}) \Phi^j + \Delta_{rr} \frac{\partial \Phi^j}{\partial \tilde{r}} + \frac{\Delta_{r\theta}}{\tilde{r}} \frac{\partial \Phi^j}{\partial \tilde{\theta}} \} \right) \tilde{r} d\tilde{r} d\tilde{\theta} \right] \alpha_j. \tag{4.96}\]

The integral in the bracket of the summation of eq. (4.96) gives the components \( A_{ij} \) of a \( N_n \times N_n \) matrix \( A \), which is not symmetric. Therefore,

\[
I_i = \sum_{j=1}^{N_n} A_{ij} \alpha_j. \tag{4.97}\]

Taking the integrals for all the \( N_n \) basis functions gives the vector

\[
\mathbf{I}_{\{N_n\}} = \mathbf{A} \cdot \mathbf{\alpha}. \tag{4.98}\]

where the vector \( \mathbf{\alpha} \) consists of all the \( N_n \) coefficients \( \{ \alpha_j \} \).

The Galerkin Finite Element method reduces the boundary value problem to the solution of a linear equation set, which is given by eqs. (4.94) and (4.98) as

\[
\mathbf{A} \cdot \mathbf{\alpha} = \mathbf{b}, \tag{4.99}\]

179
where the nonzero vector $b$ comes from the application of the essential boundary condition $\tilde{c}(\tilde{R}, \theta) = 1$ at the outer boundary of the domain $D$. The components of the matrix $A$ are calculated from the integral (4.96). $3 \times 3$ point Gaussian quadrature (Dahlquist and Björck, 1974) is used for the integration. The linear equation set (4.99) is solved by direct LU factorization (Dahlquist and Björck, 1974) to give the interpolation coefficient vector $\alpha$. The concentration field is calculated by eq. (4.90).

The calculations reported below are based on a mesh with $N_r = 28$ and $N_\theta = 36$, which leads to 4104 degrees-of-freedom in the representation of $c(r, \theta)$. The finite element mesh is graded radially to be finer near the dislocation core, where high concentration gradients are expected. Also, the angular spacing of the mesh is graded so that more elements appear below the dislocation glide plane where oxygen is attracted. The finite element mesh is shown in Fig. 4.13.

4.3.3. Oxygen Atmosphere Formation

The contours of equal oxygen concentration computed using the isotropic model are presented in Fig. 4.14 for a temperature $T = 1500$ K. The equiconcentration lines around an immobile $60^\circ$ dislocation are shown in Fig. 4.14a and are circles about the $y$–axis. The calculations agree well with the closed-form solution (4.89). The effect of dislocation motion on the oxygen equiconcentration lines is shown in Fig. 4.14b for $v = 5.5 \times 10^{-4}$ m/s; the details of this solution in the vicinity of the dislocation are shown in Fig. 4.14c. The distortion of the oxygen contours results from the dislocation motion, as oxygen is swept behind the moving dislocation. The field takes on the appearance of an oxygen plume attached to the dislocation and the plume becomes more elongated with increasing velocity, as demonstrated by Figs. 4.14d and 4.14e for $v = 3.0 \times 10^{-3}$ m/s. Steep concentration gradients appear at the ‘leading edge’ of the dislocation as the bulk concentration is convected closer to the core.

The oxygen field's computed using the anisotropic model are shown in Fig. 4.15 for $T = 1500$ K. The anisotropy causes asymmetry in the concentration field about the $y$–axis, even when the dislocation is immobile ($v = 0$), as shown in Fig. 4.15a. Instead, the
Figure 4.13. Finite element mesh used for the numerical computations of the oxygen field around 60° dislocations.
Figure 4.14. Oxygen distribution around a 60° dislocation in silicon computed using the model of isotropic migration, for $T = 1500$ K, and the dislocation velocities of (a) 0, (b)-(c) $5.5 \times 10^{-4}$ m/s, and (d)-(e) $3.0 \times 10^{-3}$ m/s. The contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 and 10.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$, and (d) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$. 
**Figure 4.15.** Oxygen distribution around a 60° dislocation computed using the model of anisotropic migration, for $T = 1500$ K, and the dislocation velocities of (a) 0, (b)-(c) $3.0 \times 10^{-4}$ m/s, and (d)-(e) $3.0 \times 10^{-3}$ m/s. The contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0 and 10.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 and 7.5 of $c_0$, and (d) 1.1, 1.2, 1.5, 2.0, 3.0, 4.0, 6.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0 and 20.0 of $c_0$. 

183
atmosphere is symmetric about the crystallographic axis \( x_2 \), as expected from the symmetry of \( \hat{u}_d \) and \( \Delta \) in eqs. (4.86). The oxygen concentration contours around a moving dislocation are shown in Figs. 4.15b and 4.15c for \( v = 3.0 \times 10^{-4} \) m/s. The distortion of the field from the contours for an immobile dislocation is caused by dislocation motion and is more severe at higher dislocation velocities, as shown in Figs. 4.15d and 4.15e for a dislocation moving with velocity \( v = 3.0 \times 10^{-3} \) m/s.

Similar calculations of the oxygen concentration field around a 60° dislocation have been carried out using both the isotropic and the anisotropic diffusion models for a range of dislocation velocities and temperatures of 1323, 1200, 1173, 1123, 1073, 920 and 600 K. The qualitative features of the oxygen concentration field for these calculations are similar to those presented in Figs. 4.14 and 4.15. Systematic experimental studies of the effects of oxygen on dislocation mobility in silicon (Imai and Sumino, 1983) and dynamic deformation of silicon (Yonenaga et al., 1984) have been carried out at the temperatures of 920 K and 1073 K. Direct comparison of our modeling results with these sets of experimental data motivates the detailed study of the oxygen concentration field at these temperatures. The formation of oxygen atmospheres and the evolution of the shape of these atmospheres with increasing dislocation velocity is presented at \( T = 920 \) K and \( T = 1073 \) K in Figures 4.16 and 4.17, respectively, using the constitutive model of anisotropic migration. The comparison with the experimental data is carried out in Chapter 5 after the development of constitutive models for dislocation drag and lattice hardening due to oxygen.

4.4. Summary and Discussion

The analysis of oxygen migration around dislocations in silicon is based on a combination of lattice-scale analysis, constitutive modeling and continuum modeling. The jump process of interstitial impurities is studied in the lattice scale and is combined with the moment analysis of the dilute impurity distribution based on the macroscopic transport equation to provide the basis for constitutive modeling of impurity migration under the action of a force field. Introduction of the structural details of the silicon lattice perpendicular to the glide plane of 60° glide dislocations transforms a general set of constitutive
Figure 4.16. Oxygen distribution around a $60^\circ$ dislocation for $T = 920$ K and dislocation velocities of (a) 0.0 m/s, (b) $2.5 \times 10^{-9}$ m/s and (c) $3.0 \times 10^{-8}$ m/s. Contours are shown of (a) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, 20.0 and 30.0 of $c_0$, (b) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0 and 20.0 of $c_0$, and (c) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 20.0 and 40.0 of $c_0$, respectively.
Figure 4.17. Oxygen distribution around a 60° dislocation for $T = 1073$ K and dislocation velocities of (a) 0, (b) $1.8 \times 10^{-7}$ m/s, and (c) $2.5 \times 10^{-6}$ m/s. The contours shown are (a) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0, 20.0, and 24.0 of $c_o$, (b) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, and 15.0 of $c_o$, and (c) 1.1, 1.2, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, 20.0, 25.0, 30.0, and 35.0 of $c_o$. 

186
equations to a constitutive model for oxygen transport in the stress field created in the lattice by a 60° dislocation. Incorporation of the constitutive model to the macroscopic equation for oxygen transport predicts the formation of oxygen atmospheres around dislocations and the evolution of the shape of these atmospheres with increasing dislocation speed.

The lattice-scale analysis was limited to the migration of interstitial impurities. The atomic jump process is more complicated for substitutional impurities and a rigorous analysis requires knowledge of the concentration of vacancies and self-interstitials in the lattice. Correlation factors are also calculated for such cases (Manning, 1968; Kehr and Binder, 1984). The constitutive equations derived in Section 4.1 also are expected to be valid for intrinsic point defects and substitutional impurities. In the case of intrinsic point defects rigorous analysis of the migration process can be carried out using Monte Carlo or molecular dynamics simulations following the Lagrangian formulation of Parrinello and Rahman (1980, 1981). However, the results are limited by the accuracy of the known interatomic potentials which is crucial for the calculation of activation energy barriers. In spite of the recent progress in the atomic modeling of silica (Tsuneyuki et al., 1988 & 1990) no empirical interatomic potential exists for a satisfactory description of dilute oxygen in silicon. A systematic analysis of the migration of vacancies and self-interstitials in silicon is presented in Chapter 6 based on atomistic simulations using the interatomic potential of Stillinger and Weber (1985).

Numerical simulation of oxygen cloud formation provides the starting point for analysis of dislocation dynamics in silicon in the presence of high oxygen concentrations. Calculation of the impurity distribution around dislocations is a first step in a systematic analysis that aims in the quantitative prediction of impurity effects on dislocation dynamics starting from the microscopic details of impurity migration and clustering under the influence of the dislocation field in the crystal. This analysis will be completed in Chapter 5 in combination with dislocation dynamics in the high-purity semiconductor and constitutive modeling of the effects of oxygen on the dynamics of plastic deformation of the crystal.
5. EFFECTS OF INTERSTITIAL OXYGEN ON THE MECHANICAL STRENGTH OF SILICON

Control of the oxygen concentration in silicon crystals (Brown, 1988) and dissolution of oxygen precipitates by preannealing of silicon wafers at high temperatures (Yonenaga et al., 1984) prevents the warpage of wafers during thermal cycling (Sumino, 1987a). Imai and Sumino (1983) demonstrated that the motion of 60° dislocations in silicon crystals grown by the Czochralski method is retarded by the presence of oxygen as compared to the motion of dislocations in high-purity silicon grown by the floating zone method. Moreover, they noted that dislocation motion stopped for applied stresses lower than a critical value, defined as the locking stress. In another series of experiments Sumino and Imai (1983) showed that immobile dislocations become mobile only when the applied stress exceeds a second critical value known as the unlocking stress. For stresses above the unlocking stress the dislocation velocity increases suddenly to a finite value. This retardation and immobilization of dislocations due to oxygen is believed to be responsible for the improvement of the mechanical strength of silicon at low deformations (Yonenaga et al., 1984). The experimental evidence indicates that the dynamics of dislocations in silicon with high oxygen concentrations is not described by the simple velocity equation (2.5) and motivates the systematic quantitative modeling of the effects of oxygen on the deformation dynamics of silicon crystals.

The development of constitutive theories that describe the effects of impurities on dislocation dynamics and on continuum mechanical properties of the material is a difficult task. This difficulty arises from the need to link atomistic understanding of the role of impurities in the lattice to micromechanical models for dislocations and impurities with macroscopic models of dislocation propagation and multiplication. The results of Chapter 4 for the oxygen concentration field around dislocations are combined in this chapter with models of dislocation dynamics and multiplication to predict the deformation behavior and mechanical properties of oxygen-hardened silicon.

A model for the resistance to dislocation motion due to oxygen is developed in Section 5.1 and is used to analyze the locking and unlocking of dislocations as a result of oxygen
gettering to the core of a dislocation. Constitutive equations for the effective stress and dislocation multiplication in oxygen-hardened silicon are developed in Section 5.2. The constitutive theories are used in Section 5.3 for the simulation of the dynamics of plastic deformation in silicon; these models quantitatively predict the effects of oxygen on the mechanical strength of the crystal.

5.1. Constitutive Modeling of Dislocation Drag. Dislocation Locking and Unlocking

The effects of oxygen on dislocation motion in silicon are expressed by the dependence of the velocity of isolated dislocations on the applied stress \( \tau \) in the crystal. This dependence was measured experimentally (Imai and Sumino, 1983; Sumino, 1987a) as a function of temperature and oxygen concentration in the crystal. The objective of constitutive modeling of the effects of oxygen on dislocation mobility is the prediction of the dependence \( v = v(\tau; T, c_o) \), where \( c_o \) is the typical oxygen level in the crystal.

The applied stress \( \tau \) is the resolved shear stress on the slip system associated with the isolated gliding dislocation. This stress is required to keep the dislocation moving at a constant speed \( v \) by counterbalancing the total drag stress that acts on the moving dislocation. This dislocation drag expresses the lattice resistance to dislocation motion which was briefly described in Section 2.1. The presence of solute significantly affects the resistance to dislocation motion leading to acceleration or retardation of the dislocation mobility as described in Section 2.3. Constitutive modeling of the effects of impurities on dislocation motion aims at predicting the contributions to dislocation drag caused by the interactions of mobile dislocations with dissolved impurity atoms or impurity precipitates in the crystal.

The analysis of this section focuses on the study of the effects of oxygen on dislocation mobility in silicon based on a constitutive equation for the total drag stress on the gliding dislocation. The total drag stress \( \tau_d \) is expressed by the linear superposition of three components

\[
\tau_d(v) = \tau_{hp}(v) + \tau_{atm}(v) + \tau_c(v),
\] (5.1)
where $\tau_{hp}$ is the drag stress inherent in dislocation motion in the high-purity material, $\tau_{atm}$ is the drag stress caused by the elastic interaction between the dislocation and the oxygen atmosphere that surrounds the dislocation core and $\tau_c$ is the drag stress due to the presence of oxygen in the dislocation core. This analysis neglects the formation of oxygen clusters in the bulk crystal, which must be present when the solubility limit for oxygen in the lattice is exceeded. The validity of the assumption that the influence of these clusters as nuclei for oxygen precipitation outside the core and as contributors to the elastic stress in the crystal is negligible was discussed in Section 4.3.1. The scalar form of the decomposition of eq. (5.1) is completely justified for a specific slip system defined by the dislocation glide plane and the glide direction. Calculation of the individual contributions to dislocation drag and inversion of eq. (5.1) gives the dependence of the dislocation velocity on the total drag stress as a function of temperature and oxygen concentration.

The phenomenological constitutive equation for the drag stress on 60° dislocations in high-purity silicon is given in Section 5.1.1. Calculation of the drag stress component that surrounds the dislocation core is presented in detail in Section 5.1.2 based on our computations of Section 4.3 for the interstitial oxygen distribution around 60° dislocations. Calculation of the drag stress component due to oxygen inside the core is based on a combination of a model for oxygen precipitate formation in dislocations with the numerical results for the oxygen distribution outside the core and is presented in Section 5.1.3. The model of dislocation drag is used to explain the locking and unlocking of dislocations due to oxygen gettering in Czochralski grown silicon in Section 5.1.4. A constitutive equation is developed for the unlocking stress as a function of temperature and oxygen concentration to be used for the modeling of the dynamic deformation of silicon.

5.1.1. Dislocation Drag in High-Purity Silicon

The velocity of isolated dislocations in the diamond/zincblende structure is modeled by equation (2.5)

$$v = B_0 \exp\left(-\frac{Q}{kT}\right) \tau^m,$$

where the phenomenological constants $(B_0, Q, m)$ depend in general on the dislocation
type and the slip system. It was mentioned in Section 3.1 that dislocation speed is equal for both 60° dislocations and screw dislocations in silicon. The validity of eq. (2.5) at low applied stresses has been discussed by Alexander (1986). Imai and Sumino (1983) have reported the correlation

\[ v = 1 \times 10^{-2} \exp\left(-\frac{2.20 \text{eV}}{kT}\right) \tau \quad \text{m/s}, \]  

(5.2)

for the velocity of 60° dislocations in high-purity silicon grown by the floating zone method. Equation (5.2) expresses the lattice resistance inherent in dislocation motion in the material of high purity because a nonzero applied stress is always required for dislocation motion. Comparison of equation (5.2) with equation (2.5) gives an exponent \( m = 1 \) for 60° dislocations in high-purity silicon.

Equation (5.2) is inverted to yield the applied stress \( \tau(v) \) as a function of temperature and dislocation velocity. This applied stress is equal to the component \( \tau_{hp}(v) \) in the decomposition of equation (5.1). The drag stress inherent in dislocation motion in the high-purity material is

\[ \tau_{hp}(v) = \frac{1}{B_0} \exp\left(\frac{Q}{kT}\right) v, \]  

(5.3)

where \( Q = 2.20 \text{ eV} \) and \( B_0 = 1.0 \times 10^{-2} \text{ m}^{-1} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1} \) according to the correlation (5.2).

### 5.1.2. Dislocation Drag from Interaction with Oxygen Atmospheres

The elastic interaction potential (4.69) determines the force exerted by the dislocation on the oxygen atoms. An equal and opposite force is exerted by the oxygen atoms on the dislocation. Therefore, as the dislocation moves, this force provides a mechanism for dislocation drag. The drag force \( \mathbf{F}_D \) is computed from the oxygen concentration field as \( \mathbf{F}_D = -F_D \hat{x} \), where \( F_D \) is the magnitude of the force and the direction of the drag force is set by the direction of the dislocation motion.

The drag force caused by the oxygen distribution according to the isotropic transport model of Section 4.3.1.2 is computed in the cartesian coordinate system \((x, y)\) by integrating the component \( -F_x \) of the force \( -\mathbf{F} = -F_x \hat{x} - F_y \hat{y} \) due to each impurity atom.
throughout the domain. The component $-F_x$ is opposed to the dislocation motion and determines the drag force per unit length (N/cm) of the dislocation $F_D$ as

$$F_D = \int_D c(r, \theta)F_x(r, \theta) \ r dr d\theta, \tag{5.4a}$$

or in terms of the dimensionless variables defined in Section 4.3.1.1

$$F_D = c_0 a^2 \int_1^R \int_0^{2\pi} \tilde{c}(\tilde{r}, \theta)F_x(\tilde{r}, \theta) \ \tilde{r} d\tilde{r} d\theta. \tag{5.4b}$$

This equation was initially developed by Cottrell and Jawson (1949) to calculate drag stresses using an asymptotic approximation for the concentration field that is valid for $Pe \ll 1$, where the dislocation Peclet number $Pe$ scales the ratio of convection due to dislocation motion with solute diffusion. The contribution to the drag force by oxygen atoms with $r > R$ in the computational domain of Fig. 4.12 is small because of the relatively weak oxygen-dislocation interaction far away from the core and has been neglected in eq. (5.4). The drag force (5.4) is zero for the immobile dislocation because the oxygen distribution is symmetric about the $\hat{y}$–axis.

The drag force predicted using the anisotropic model of oxygen diffusion is expressed in terms of the crystallographic axes $(x_1, x_2)$, where the force on the dislocation caused by each of the impurity atoms is written as $\mathbf{F} = F_1 \hat{x} + F_2 \hat{x}_2$. The component $-F_1$ opposes dislocation motion and determines the drag force $F_D$. Expressing the force component $F_1$ in terms of the cartesian components of eq. (4.71a) gives

$$F_D = \int_D c(r, \theta)[F_x(r, \theta) + \frac{\sqrt{3}}{4} F_y(r, \theta)] \ r dr d\theta, \tag{5.5a}$$

or

$$F_D = c_0 a^2 \int_1^R \int_0^{2\pi} \tilde{c}(\tilde{r}, \theta)[F_x(\tilde{r}, \theta) + \frac{\sqrt{3}}{4} F_y(\tilde{r}, \theta)] \ \tilde{r} d\tilde{r} d\theta. \tag{5.5b}$$

For the immobile dislocation, the oxygen distribution is symmetric about the $x_2$–axis and so the drag force is equal to zero. The drag force in both expressions (5.4) and (5.5)
is proportional to the initial impurity concentration \( c_0 \) for constant temperature \( T \) and dislocation velocity \( v \).

The drag force exerted on the dislocation by the oxygen atmosphere is calculated using both the isotropic and the anisotropic models using the computed concentration fields \( c(r, \theta) \) and Gaussian quadrature for evaluating the integrals. The dependence of \( F_D \) on the dislocation velocity \( v \) is shown in Fig. 5.1 for the anisotropic diffusion model with \( T = 1500 \) K. It exhibits a maximum value at some intermediate critical velocity \( v_c \), shown as point B on Fig. 5.1. Velocities higher than the critical velocity are associated with severe distortions of the oxygen atmosphere and further increases in the dislocation velocity lead to decreases in \( F_D \). It should be mentioned that the oxygen saturation boundary condition \( (4.81) \) guarantees that the oxygen atmosphere is always attached to the dislocation core. Therefore, the unstable dislocation motion for \( v > v_c \) does not mean that the dislocation breaks away from the oxygen atmosphere.

Cottrell (1953) approximated the critical velocity by

\[
v_c \simeq \frac{4kTD}{A_d b}.
\]  

(5.6)

The derivation of eq. (5.6) was based on the assumption that the critical dislocation velocity is equal to a characteristic drift of the solute (Cottrell, 1953). The critical dislocation velocity is computed to be \( v_c = 3.0 \times 10^{-4} \) m/s using the model for anisotropic diffusion. With the appropriate values of \( A_d = 0.5 \) eV, \( b = 3.84 \times 10^{-10} \) m, \( T = 1500 \) K and eq. (4.75) for \( D \), Cottrell's expression (5.6) yields \( v_c = 2.1 \times 10^{-4} \) m/s, in fairly good agreement with our calculations. Yoshinaga and Morozumi (1971) calculated the drag force-velocity curve \( F_D = F_D(v) \) numerically for the case of carbon gettering by edge dislocations in vanadium. They also found Cottrell's approximation to be satisfactory. The oxygen concentration distribution that has been shown in Fig. 4.15b corresponds to the critical dislocation velocity \( v_c \) for \( T = 1500 \) K. The concentration field that corresponds to point C of Fig. 5.1 has been shown in Fig. 4.15d.

It is interesting to compare the predictions of the dependence of the drag on the dislocation velocity for the isotropic and anisotropic transport models of Section 4.3.1.2.
Figure 5.1. Dislocation velocity as a function of the drag force developed by the interaction between the dislocation and the oxygen distribution; $T = 1500$ K.
Results are shown in Fig. 5.2 for temperatures in the range $600 \, \text{K} \leq T \leq 1500 \, \text{K}$. The isotropic model underestimates the drag force at low dislocation velocities and high temperatures, $920 \, \text{K} \leq T \leq 1500 \, \text{K}$, but overestimates it at high dislocation velocities. At lower temperatures, where the anisotropy of the diffusivity tensor is significant, the isotropic diffusion model overestimates the drag force by as much as 50 percent.

The temperature dependence of the maximum drag force $F_{D,\text{max}}$ which corresponds to the critical dislocation velocity $v_c$ is shown as an Arrhenius plot in Fig. 5.3. The Arrhenius behavior of the force is expected because of the Boltzmann factors which appear in the expressions of the diffusivity and the drift velocity. The maximum drag force decreases with increasing temperature. This observation was made initially by Yoshinaga and Morozumi (1971), but calculations of the Arrhenius behavior were not presented. The comparison between the isotropic and the anisotropic diffusivity models in Fig. 5.3 shows that the isotropic model overestimates the maximum drag force at low temperatures, but underestimates it at $T > 1200 \, \text{K}$.

The temperature dependence of the critical dislocation velocity $v_c$ is shown in Fig. 5.4 for the isotropic and anisotropic transport models and is compared with the prediction of Cottrell, eq. (5.6). Equation (5.6) underestimates $v_c$; this observation was made first by Yoshinaga and Morozumi (1971).

The stress field $\mathbf{g}$ that creates a force $\mathbf{F}$ on a dislocation line is related to $\mathbf{F}$ by the Peach-Köhler equation (Hirth and Lothe, 1982; Nabarro, 1987)

$$\frac{\mathbf{F}}{L} = (\mathbf{b} \cdot \mathbf{g}) \times \mathbf{t}_d,$$  \hspace{1cm} (2.3)

where $L$ is the length of the dislocation line, $\mathbf{b}$ is the Burgers vector and $\mathbf{t}_d$ is the unit vector tangent to the dislocation line. If the stress tensor is caused by dragging the impurity field and the resolved shear stress component $\sigma_{xy}$ is nonzero and equal to $\tau_{\text{atm}}$, eq. (2.3) takes the simple form

$$F_D = b \tau_{\text{atm}},$$  \hspace{1cm} (5.7)

so that the functional dependence of the drag stress $\tau_{\text{atm}}$ used in equation (5.1) is identical to the drag force $F_D$. 

195
Figure 5.2. Comparison of the predictions of the isotropic diffusivity model (○) and the anisotropic diffusivity model (●) for the dependence of the dislocation velocity on the drag force; Results are shown for the temperatures (a) $T = 1500$ K, (b) $T = 1200$ K, (c) $T = 920$ K and (d) $T = 600$ K.
Figure 5.3. Arrhenius plot of the temperature dependence of the maximum value of the dislocation drag force for the isotropic (○) and the anisotropic (□) diffusivity models.
Figure 5.4. Temperature dependence of the critical dislocation velocity predicted from the isotropic (○) and the anisotropic (□) diffusivity models. Cottrell's approximation (---) also is shown.
It is interesting to calculate the dislocation drag stress $\tau_d$ in the presence of oxygen assuming that the contribution $\tau_c(v)$ in eq. (5.1) is negligible. Therefore, the effect of oxygen is limited to the drag stress due to the atmosphere $\tau_{atm}(v)$, which is calculated by eq. (5.7) and is superposed to the drag in the high-purity material given by eq. (5.3). The result is plotted as Fig. 5.5. for an oxygen concentration $c_0 = 1 \times 10^{18}$ cm$^{-3}$ and $T = 920$ K. The increased drag on the dislocation caused by the deformation of the impurity atmosphere retards dislocation motion for a specific oxygen concentration. The magnitude of the retardation increases with increasing $c_0$ according to eq. (23b) and is demonstrated for a range of oxygen concentrations in Fig. 5.6. At low stresses and low velocities, the symmetry of the oxygen atmosphere makes this component of the drag unimportant and the relationship $v = v(\tau_d)$ parallels the result for pure silicon.

The predictions for the dependence of the dislocation velocity on the applied stress computed by neglecting the contribution to the drag stress due to oxygen inside the core are compared in Fig. 5.7 directly to the measurements of Imai and Sumino (1983) for oxygen concentrations of $7.4 \times 10^{17}$ and $9.2 \times 10^{17}$ cm$^{-3}$. In their experiments, Imai and Sumino introduced dislocations into a Czochralski grown sample by making a Knoop indentation (Imai and Sumino, 1983) and placing the sample in tension at 15-20 MPa at temperatures between 823 and 923 K until specifically sized dislocation loops were observed. The stress level was slowly decreased until dislocation motion was no longer observed. The results of our calculations are in qualitative agreement at high stress values and demonstrate the correct form of the retardation of the dislocation motion caused by the oxygen atmosphere. However, the predictions do not capture the locking stress at which the dislocations cease to move. This is not surprising since both components $\tau_{atm}(v)$ and $\tau_{hp}(v)$, tend to zero as the velocity decreases and emphasize the importance of accurate modeling of the effects of oxygen inside the dislocation core on dislocation dynamics.

5.1.3. Dislocation Drag from Oxygen Precipitates

Oxygen precipitation is taken into account only within the dislocation core and gives rise to the drag stress $\tau_c$. At steady state the amount of oxygen in the core is constant,
Figure 5.5. Dependence of the dislocation velocity (---) on the applied stress for oxygen concentration $c_0 = 1.0 \times 10^{18}$ cm$^{-3}$ at $T = 920$ K. The predicted dependence of the dislocation drag stress on the dislocation velocity (-----) and the stress dependence of the dislocation velocity for pure silicon (- - -) also are shown.
Figure 5.6. Effect of bulk oxygen concentration on the dependence of dislocation velocity on the applied stress for $T = 920$ K.
Figure 5.7. Comparison of the predicted dependence of the dislocation velocity on the applied stress neglecting the drag stress due to oxygen inside the core with the experimental data of Imai and Sumino; $T = 920$ K; (a) $c_0 = 7.4 \times 10^{17}$ cm$^{-3}$ and (b) $c_0 = 9.2 \times 10^{17}$ cm$^{-3}$.
but individual atoms are continuously exchanged between the dislocation core and the lattice. The no-flux boundary condition (4.81) at \( r = r_0 \), which was used for the calculation of the oxygen distribution implies that the oxygen concentration in the core, \( c_{\text{core}} \), is equal to the oxygen concentration on the boundary between the dislocation core and the lattice; a similar boundary condition was introduced by Bullough and Newman (1970). Computing the oxygen concentration in the core as the polar average of \( c(r_0, \theta) \) gives the number of oxygen atoms there as \( N_c = c_{\text{core}} \pi r_0^2 \). In our calculations, the oxygen concentration \( c_{\text{core}} \) varies as a function of the dislocation velocity from 4.1 to 38.7 of \( c_o \) at \( T = 920 \) K.

The drag stress \( \tau_c \) is caused by the consumption of the energy needed to dissociate oxygen from the core. A model for \( \tau_c \) is developed by defining the energy difference \( \Delta E^o \) between an oxygen atom in the core and an isolated interstitial oxygen atom in the lattice adjacent to it. The starting point in the analysis is the understanding of the structure and the energetics of oxygen inside the dislocation core. Oxygen precipitates inside the core are modelled as quasi-one-dimensional chain-like clusters, as motivated by the recent \textit{ab initio} calculations of Needels et al. (1991) for oxygen aggregates in bulk silicon. According to these results, oxygen atoms preferentially form staggered clusters with large binding energies in bridging configurations of the form -O-Si-O-Si-O-. In particular, each new oxygen atom is bound to the chain with a binding energy \( E_b = 1.0 \) eV with respect to interstitial oxygen. The structure of the stable quasi-one dimensional oxygen cluster is shown in Fig. 5.8.

The same configuration is assumed for oxygen atoms inside the core so that the predicted binding energy is \( E_b = 1.0 \) eV (Needels et al., 1991) with respect to isolated interstitial oxygen in a bridging configuration. Therefore, \( \Delta E^o \) is defined as

\[
\Delta E^o \equiv E_b - E_i, \tag{5.8}
\]

where \( E_i = 0.5 \) eV is the interaction energy between an oxygen atom at an interstitial site in the lattice adjacent to the core and the dislocation (Sumino and Imai, 1983). Sumino and Imai (1983) presented a phenomenological model for dislocation unlocking in silicon as a thermally activated process with an activation energy of 3.0 eV. This value can be derived from our model by adding the energy difference \( \Delta E^o = 0.5 \) eV to the activation
Figure 5.8. Structure of staggered cluster of oxygen in bulk silicon (Needels et al., 1991).
energy (2.5 eV) for oxygen migration in silicon (Inoue et al., 1987; Needels et al., 1991).

The contribution to the drag stress $\tau_c$ is computed using the Peach-Koehler equation (2.3) as

$$\tau_c = \frac{F_c}{b} = \frac{1}{b} \frac{N_c \Delta E^o}{a},$$  \hspace{1cm} (5.9)

where the energy gradient $\Delta E^o/a = \Delta E^o/r_0$ is the force required to dissociate one oxygen atom from the core precipitate and put it at an interstitial site next to the core. The dependence of $\tau_c$ on the dislocation velocity $v$ arises simply from the variation of $c_{\text{core}}$ with $v$, and is shown in Fig. 5.9 for $c_o = 9.2 \times 10^{17}$ cm$^{-3}$ and $T = 920$ K.

The result for the drag stress $\tau_{\text{atm}}(v)$ due to the interaction of the dislocation with the oxygen atmosphere also is shown in Fig. 5.9 for comparison with $\tau_c(v)$. Refinement of the finite-element mesh permits the calculation of the oxygen concentration field for a wide range of dislocation velocities. The comparison of the two contributions to dislocation drag leads to two interesting conclusions. First, the magnitude of the contribution to the drag stress due to oxygen precipitates in the core is much higher than the contribution due to the atmosphere of dissolved interstitial oxygen; typically by one to two orders-of-magnitude. Second, the drag stress due to the oxygen precipitates initially decreases with increasing dislocation velocities. A minimum in the drag stress is reached at a critical dislocation velocity and further increase in the velocity leads to increasing drag stress. Therefore, independent examination of the two mechanisms which contribute to dislocation drag due to the presence of oxygen in the crystal leads to opposite conclusions for the stability of dislocation motion. This means that the complete picture of dislocation dynamics provided by equation (5.1) will be different from the simple picture provided by Cotrell’s approximation.

5.1.4. Modeling of Dislocation Locking and Unlocking

The complete form of the constitutive equation (5.1) is used for the calculation of the resistance to dislocation motion after the investigation of the individual contributions to dislocation drag. The predictions from eq. (5.1) of the relationship $v = v(\tau; T, c_o)$ are directly compared in Fig. 5.10 with the experimental data of Imai and Sumino (1983) for
Figure 5.9. Relationship between dislocation velocity and drag stress due to (1) the elastic interaction between the oxygen atmosphere in the lattice and the dislocation ($\tau_{\text{atm}}$) and (2) oxygen precipitates inside the core ($\tau_c$). The calculations are for a silicon crystal with oxygen concentration $c_o = 9.2 \times 10^{17} \text{ cm}^{-3}$ at $T = 920 \text{ K}$. 
Figure 5.10. Comparison between the model predictions for the dependence of the 60° dislocation velocity on the applied stress in silicon crystals with oxygen concentrations (a) $c_0 = 2.5 \times 10^{17}$ cm$^{-3}$, (b) $c_0 = 7.4 \times 10^{17}$ cm$^{-3}$ and (c) $c_0 = 9.2 \times 10^{17}$ cm$^{-3}$ at $T = 920$ K (---) and the experimental data of Imai and Sumino ($\triangle$, $\bigcirc$, $\square$). The dashed line gives the result in high purity silicon.
Czochralski grown silicon with oxygen concentrations of 2.5, 7.4 and \(9.2 \times 10^{17}\) cm\(^{-3}\). The agreement between the experimental data and eq. (5.1) is very good. Interestingly, the model predicts three regimes of dislocation motion. At low applied stresses no dislocation motion is predicted and the dislocation is locked in place. At a critical applied stress, the unlocking stress \(\tau_u\), the dislocation begins to move; however, initially the dislocation velocity increases with decreasing stress and these steady states are unstable. At a second critical stress, the locking stress \(\tau_l\), the dislocation velocity increases with applied stress and the steady-state motion is stable. Both the locking and the unlocking stress are proportional to the oxygen concentration \(c_o\).

The locking stresses for the three crystals used in the experiments were reported to be 1.8, 2.7 and 8.0 MPa (Imai and Sumino, 1983), compared to the predictions of 0.8, 1.8 and 2.2 MPa by eq. (5.1). Therefore, equation (5.1) underpredicts the increase of the locking stress with increasing oxygen concentration. The unlocking stress has been reported by Sumino and Imai (1983) to be proportional to the amount of oxygen in the core, which is proportional to \(c_o\) in our model. However, the experiments of Sumino and Imai, were not in steady state, because \(N_c\) depended on the duration of sample ageing after releasing the applied stress. Oxygen atoms are gettered to the core of the immobile dislocation during ageing at zero applied stress until saturation occurs. When the stress is applied again, the dislocation does not reach its final velocity immediately and equilibrium at the core/matrix interface is not necessarily established. Therefore, direct comparison of these experiments with the predictions of our steady state analysis is questionable.

The important conclusion of our analysis is that modeling dislocation drag by eq. (5.1) predicts hysteresis in the dislocation velocity with applied stress in silicon crystals with high oxygen concentrations. This hysteresis explains the difference in the unlocking and locking stresses for dislocation motion.

The dependence of the dislocation velocity on the applied stress is presented in Fig. 5.11 for five oxygen concentrations at \(T = 920\) and 1173 K, respectively. The hysteresis is more pronounced at the lower temperature and at the highest oxygen content in the material. For higher temperatures and lower oxygen concentrations the hysteresis disappears and the values of the locking and unlocking stresses coincide. Another interesting
Figure 5.11. Dislocation velocity as function of the applied stress for various concentrations of interstitial oxygen at (a) $T = 647^\circ$C and (b) $T = 900^\circ$C. The dashed line gives $v = v(\tau)$ for high purity material.
conclusion of the numerical simulations is that dislocation velocity is proportional to the applied stress for applied stresses higher than 3-4 MPa; this prediction is in agreement with the experimental data reviewed by Alexander and Haasen (1968).

The unlocking stress is a very important quantity for the analysis of the initial stage of deformation of silicon. This critical stress is proportional to the characteristic oxygen content $c_o$ of the material, as is implied by the proportionality to $c_o$ of $\tau_{atm}$ and $\tau_c$ in eq. (5.1). The computed dependence of the unlocking stress on oxygen concentration and temperature is presented in Figs. 5.12 and 5.13, respectively. These numerical results for the unlocking stress are modeled accurately as a thermally activated process by the correlation

$$\tau_u (Pa) \simeq 2.2 \times 10^{-13} c_o \ (cm^{-3}) \ \exp \left( \frac{0.215 \ eV}{kT} \right). \quad (5.10)$$

### 5.2. Constitutive Model for the Initial Stages of Deformation

The results for the dependence of the dislocation velocity on the applied stress presented in Section 5.1 describe the response of a single dislocation under stress in the material characterized by the presence of interstitial oxygen atoms at a characteristic level $c_o$. To understand the deformation behavior of the solute-hardened material the applied stress must be replaced by the appropriate expression for an effective resolved shear stress to account for the interactions between dislocations and between dislocations and impurity atmospheres formed around dislocation cores. The effective stress is an important quantity for the quantitative description of the deformation behavior, especially at the high dislocation densities which occur as a result of dislocation multiplication during the deformation process. We adopt the macroscopic constitutive model of Alexander and Haasen (1968) presented in Section 3.2.1 for the description of dislocation dynamics in high-purity silicon and extend it to incorporate solid solution hardening due to an oxygen impurity.

#### 5.2.1. Constitutive Model for the Effective Stress

The concept of the effective stress was originally introduced by Haasen (1962). An expression for a back stress due to the interaction between dislocations was formulated in
Figure 5.12. Critical stress $\tau_u$ for dislocation unlocking as a function of oxygen concentration in silicon as a function of temperature.
Figure 5.13. Temperature dependence of the unlocking stress as a function of oxygen concentration.
Section 3.2.1 as

\[ \tau_b = A \sqrt{N}, \]  

(5.11)

where \( A \) is the strain hardening factor (Alexander and Haasen, 1968), which depends on the magnitude of the Burgers vector of the dislocation \( b \) and on the elastic constants of the material and \( N \) is the dislocation density per unit area. The effective stress is defined as the difference between the applied stress \( \tau \) in the solid and the back stress. If the back stress exceeds the applied stress, the effective stress is zero. Therefore, the back stress expresses the resistance to dislocation motion caused by the interaction between the distinguished dislocation and others in a regular array of parallel dislocations which cause the effective stress:

\[ \tau_{\text{eff}} \equiv \begin{cases} 
\tau - \tau_b, & \tau > \tau_b, \\
0, & \tau \leq \tau_b. 
\end{cases} \]  

(5.12)

The expression for the effective stress is extended to the solute-hardened crystal by introducing an internal stress as the sum of three components:

\[ \tau_i = \tau_u + \tau_b + \tau_{i, d-c}, \]  

(5.13)

where \( \tau_u \) is the unlocking stress described in Section 5.1.4, \( \tau_b \) is the back stress and \( \tau_{i, d-c} \) is an internal stress due to the interactions between dislocations and impurity clouds formed around other dislocations. The unlocking stress describes the interaction between a dislocation and its own impurity cloud and it is included in eq. (5.13) as a link between the dynamics of isolated dislocations and dislocation networks in solute-hardened materials. For a solute-hardened material, the effective stress is given by

\[ \tau_{\text{eff}} \equiv \begin{cases} 
\tau - \tau_i, & \tau > \tau_i, \\
0, & \tau \leq \tau_i. 
\end{cases} \]  

(5.14)

Modeling the dislocation field as a regular periodic array of dislocations gives the back stress as eq. (5.11). In particular, for the case of Czochralski grown silicon, the dislocation field has been modeled as a regular periodic array of 60\(^\circ\) dislocations. This approach was partially motivated by the very close velocities of 60\(^\circ\) and screw dislocations at some
specific temperature (Imai and Sumino, 1983; Alexander, 1986). Suezawa et al. (1979) expressed
the strain hardening factor $A$ in eq. (5.11) by the phenomenological equation

$$ A = \frac{Gb}{\gamma}, \quad (5.15) $$

where $G$ is the shear modulus of the material and $\gamma = 3.3$ is a fitting parameter for the experimental data at $T = 1073$ K.

We express the back stress as a stress field experienced by a 60° dislocation due to another 60° dislocation at a distance $r = 1/\sqrt{N}$ as (Hirth and Lothe, 1982)

$$ \tau_b = \frac{Gb\sin \beta}{2\pi(1 - \nu)} \sqrt{N}, \quad (5.16) $$

where $\beta = 60°$ and $\nu$ is the Poisson ratio of the material. The shear modulus $G$ is a function of both temperature and solute concentration. A correlation reported by Dillon et al. (1987) for the Young modulus $E$ of silicon is used to express the temperature dependence of this parameter as

$$ E \text{ (Pa)} = 1.7 \times 10^{11} - 2.771 \times 10^4 T^2, \quad (5.17) $$

where $E \equiv 2G(1 + \nu)$. The value $\nu = 0.278$ was used for the Poisson ratio (Lambropoulos, 1987). The correlation (5.17) satisfactorily describes experimental data for the elastic constants of silicon (Landolt-Böhnstein, 1979) and recent results of molecular dynamics simulations (Kluge et al., 1986) and is adopted because of its simplicity.

The dependence of the shear modulus on the solute concentration is modeled using the theory described by Eshelby (1976) for the interaction between point defects. According to Eshelby’s results

$$ \frac{1}{G} \frac{dG}{dx} \simeq +2, \quad (5.18) $$

for rigid inclusions, typically interstitial atoms, where $x$ is the atomic fraction of the rigid inclusions. These inclusions are taken to be oxygen interstitial atoms for the case of Czochralski grown silicon. Equation (5.18) gives

$$ G \simeq G_0 \exp(2x), \quad (5.19) $$
where $G_0$ is the shear modulus of the high purity material. For the maximum oxygen concentration used in our calculations, $c_o = 9.0 \times 10^{17} \text{ cm}^{-3} = 18.0 \text{ ppma}$, eq. (5.19) gives

$$(G - G_0)/G_0 \simeq 3.6 \times 10^{-5},$$

which implies that the solute effect on the shear modulus of the material is negligible.

The internal stress $\tau_{i,d-c}$ in eq. (5.13) is modeled by the stress field caused by the elastic interaction between the $60^\circ$ dislocation and the oxygen cloud around another $60^\circ$ dislocation at a distance $r = 1/\sqrt{N}$, where the dominant dislocation-solute interaction is described by the elastic interaction potential of eq. (4.69). The interaction between a dislocation and a neighboring oxygen cloud is modeled by superposing the interactions due to each oxygen atom of the cloud. The oxygen atoms inside the dislocation core are excluded from the calculation because they supply the dominant contribution to the unlocking stress, which is modeled as $\tau_u$.

The calculation of $\tau_{i,d-c}$ in eq. (5.13) is based on Cottrell’s approximate analysis for an impurity cloud (Cottrell, 1953) and our numerical simulations of atmosphere formation. The geometrical details are given in Fig. 5.14. According to Cottrell’s approximation, (Cottrell, 1953) the radius of the solute atmosphere around a $60^\circ$ dislocation is approximated as

$$l_{\text{atm}} \simeq \frac{\sqrt{3}}{2} \frac{A_d b}{kT}. \quad (5.20)$$

Then the number of solute atoms $N_s$ in the cloud per unit length of dislocation is simply written as

$$N_s = \pi l_{\text{atm}}^2 \tilde{c}, \quad (5.21)$$

where

$$\tilde{c} = \alpha c_o \quad (5.22)$$

is the average solute concentration in the cloud. The numerical factor $\alpha > 1$ in eq. (5.22) depends on the dislocation velocity and temperature and is determined from the numerical simulations of the oxygen distribution. The calculation of $\tau_{i,d-c}$ is simplified by comparing $l_{\text{atm}}$ to $r = 1/\sqrt{N}$. The ratio $l_{\text{atm}}/r \ll 1$ implies that the cloud can be satisfactorily
Figure 5.14. Geometrical details for the calculation of the component of the internal stress due to the interaction between a dislocation and a neighboring oxygen cloud. The dislocation glide plane is perpendicular to the y axis.
modeled as a point inclusion with strength $N_s$ times of an individual interstitial impurity atom. Neglecting the angular dependence of the interaction potential in eq. (4.69), to be consistent with Cottrell’s approximation, gives

$$U_{\text{tot}} \simeq \frac{\sqrt{3}}{2} N_s \frac{A_d b}{r}.$$  

(5.23)

The gradient of $U_{\text{tot}}(r)$ gives the elastic interaction force. The Peach-Koehler equation (2.3) is used to transform this interaction force to the stress field $\tau_{i,d-c}$ as a function of $N$ and $c_o$. The final expression is

$$\tau_{i,d-c} = \frac{3\sqrt{3}}{8} \pi \alpha \frac{A_d^2 b^2}{(kT)^2} c_o N.$$  

(5.24)

An interesting implication of eq. (5.24) is that the internal stress due to the cloud-dislocation interaction is proportional to the product of the impurity content expressed by $c_o$ and the dislocation density $N$.

Although closed-form expressions have been derived for the components of the internal stress, some parameters in these expressions are calculated from numerical computations. A general phenomenological model is constructed for the internal stress in solute-hardened crystals, based on eqs. (5.10), (5.13), (5.16) and (5.24), as

$$\tau_i = k_c c_o + A\sqrt{N} + B c_o N.$$  

(5.25)

where $k_c$ is given by the expression for the unlocking stress, eq. (5.10), $A$ is the strain hardening factor given by eq. (5.16), and the factor $B$ is given by eq. (5.24) for the cloud-dislocation interaction. This phenomenological model includes the contributions of oxygen to the hardening of the lattice and is used in Section 5.3 for the analysis of dynamic deformation experiments.

5.2.2. Constitutive Model of Dislocation Multiplication

Dislocation multiplication due to the cross-slip mechanism has been analyzed in detail in Sections 2.2.3 and 3.2. Here the derivation is given of the kinetic expression used in the
model of Alexander and Haasen (1968). This expression is extended to incorporate the effects of solute. Johnston and Gilman (1960) and later Low and Turkalo (1962) beautifully demonstrated the geometry and described the physics of the cross-slip multiplication process in LiF and iron with 3% silicon, respectively. Both groups correctly focused their analyses on the origin of the multiplication process, which is the formation of jogs in the screw dislocation as a result of the intersection of two lines. The geometrical details of the multiplication mechanism driven by jogs of the appropriate size are shown in Fig. 5.15.

The differential increase \( dN \) in the number of dislocation lines per unit area is written as

\[
dN = N \, dn, \tag{5.26}
\]

where \( dn \) is the increase in the number of jogs which lead to dislocation multiplication by reaching the appropriate length. If the distance \( dx \) traveled by the screw dislocation moving with velocity \( v \) on the \( \{111\} \) plane is taken to be equal to the corresponding distance traveled by the 60° dislocation that intersects it, then \( dx = v \, dt \). If \( z_0 \) is the height of the formed jog at time \( t \) the increase in the number of jogs is written as

\[
dn = K \frac{dx}{z_0}, \tag{5.27}
\]

where \( K \) is a kinetic parameter. Accordingly, the number of jogs increases proportionally to the distance traveled by the dislocation and inversely proportionally to the height of the formed jog. The jog of height \( z_0 \) does not move and separates two edge dislocations formed by the slip of the original screw, as shown in Fig. 5.15. The maximum interaction between these two edge dislocations is expressed by the effective stress in the material \( \tau_{\text{eff}} \) and is related to \( z_0 \) by

\[
\tau_{\text{eff}} = \frac{Gb}{8\pi(1-\nu)} \frac{1}{z_0}. \tag{5.28}
\]

Combining eqs. (5.26)-(5.28) gives the kinetics of dislocation multiplication as

\[
\frac{dN}{dt} = K \frac{8\pi(1-\nu)}{Gb} \tau_{\text{eff}} vN, \tag{5.29}
\]
Figure 5.15. Geometrical details of the initiation of the cross-slip multiplication process due to the formation of a jog of height $z_0$ in a screw dislocation.
where $K$ is a dimensionless dislocation multiplication rate constant. Equation (5.29) was first written in this form by Alexander and Haasen (1968).

We extend eq. (5.29) to a solute-hardened crystal by writing the dislocation multiplication rate constant in the Arrhenius form

$$K = K_0 \exp\left(-\frac{\tau_u b^2 l_s}{kT}\right),$$  \hspace{1cm} (5.30)

where $K_0$ is the multiplication rate constant in the high purity material and $Q_u = \tau_u b^2 l_s$ is the activation energy for dislocation unlocking. Dislocation unlocking is the first step in the description of dislocation multiplication as a two-step thermally activated process. The second step is the formation of a jog as a result of the intersection of two dislocations. Dislocation unlocking and the consequent dislocation motion are required for dislocation intersection and jog formation.

The length $l_s$ in the expression (5.30) of the activation energy for dislocation unlocking is the equilibrium separation length between jogs of the appropriate size for multiplication in the screw dislocation line. This length $l_s$ is calculated from the value of the multiplication rate constant $K_0$ in the high purity material and is taken to be proportional to the equilibrium concentration of possible jog sites to start the dislocation multiplication process. The proportionality constant is the unit length of a dislocation line. Therefore,

$$K_0 = l_s^{-1} \exp\left(-\frac{W_f}{kT}\right),$$  \hspace{1cm} (5.31)

where $W_f$ is the energy of jog formation in the screw dislocation. For a screw dislocation $W_f$ is equal to the energy of kink formation (Hirth and Lothe, 1982) and also is equal to the activation energy for dislocation glide. The energy of kink formation has been calculated by Labusch (1965) to be equal to 2.14 eV in silicon, while the activation energy for dislocation glide measured by Yorikawa and Sumino (1978) is 2.17 eV and is in good agreement with the theoretical estimate.

We use the value of $W_f = 2.17$ eV and the value of the multiplication rate constant in high purity silicon reported by Suegawa et al. (1979) to calculate $l_s$. The value of $K \cdot 8\pi(1-\nu)/Gb = 3.1 \times 10^{-4}$ m/N is given by the previous authors as appropriate for
fitting stress-strain curves from dynamic deformation experiments at 1073 K. The length \( l_s \) is calculated to be equal to \( 1.63 \times 10^{-7} \) m. Although eq. (5.31) implies the temperature dependence of \( K_0 \), the same value of \( l_s \) is retained in the temperature range 1073 – 1273 K. There are two reasons for this decision. First, no experimental data for \( K_0 \) exist at temperatures other than 1073 K. Second \( l_s \) is in the correct order-of-magnitude, as it is implied by the experimental observations of Low and Turkalo (1962) for the separation length between jogs equal to 0.5 \( \mu \text{m} \) in iron with 3% silicon for 298 K \( \leq T \leq 323 \) K. Also, these authors observed that the minimum length of the dipole loops left in the wake of a screw dislocation with jogs of intermediate size was 0.3 \( \mu \text{m} \). Eqs. (5.29) and (5.30) provide the basis for the analysis of dislocation multiplication in the solute-hardened crystal.

5.3. Modeling of Plastic Deformation Dynamics

The results for the mobility of isolated dislocations presented in Section 5.1.4 and the constitutive models for the effective stress and dislocation multiplication in the oxygen-hardened crystal developed in Section 5.2 combined with ideal plasticity theory (Boley and Weiner, 1960) and Orowan's equation (3.13) provide a framework for the analysis of plastic deformation dynamics in silicon crystals with various oxygen concentrations. Dynamic deformation experiments are simulated in Section 5.3.1 to monitor the evolution of stress and dislocation density in the crystal during deformation. These simulations provide a clear picture of the effects of oxygen on the mechanical strength of silicon and have important implications in the thermal processing of silicon wafers. Approximate expressions for the estimation of the critical resolved shear stress are given in Section 5.3.2 and the limiting cases of small and large deformations are analyzed further in Section 5.3.3.

5.3.1. Simulation of Dynamic Deformation Experiments

Dynamic deformation experiments are carried out under constant strain rate \( \dot{\varepsilon} \) and temperature. A large amount of data on the dynamic deformation of diamond/zincblende structure materials has been obtained from the work of Haasen and coworkers in Germany (Alexander and Haasen, 1968) and Sumino and coworkers in Japan (Sumino, 1987a). Here
we focus on the dynamic deformation of Czochralski grown silicon and simulate dynamic deformation experiments carried out by Yonenaga et al. (1984).

Using fundamental equations of idealized plasticity theory (Boley and Weiner, 1960) and resolving the stress and strain tensors on the primary slip system, the applied stress is defined (Suezawa, 1979) by

$$\tau = \tau_{el} - \eta \varepsilon^p, \quad (5.32)$$

where \( \tau_{el} \) is the resolved shear stress as if the material was purely elastic, \( \varepsilon^p \) is the resolved plastic strain and \( \eta \) is the elastic constant of the composite system composed of the machine and the specimen for shear deformation. The elastic constant \( \eta \) is generally a function of \( \dot{\varepsilon} \) and \( T \) (Suezawa, 1979). The plastic response of the material under stress expressed by eq. (5.32) is linked to a micromechanical picture of plastic flow due to dislocation motion by the Orowan equation (3.13) for the plastic strain rate as

$$\frac{d \varepsilon^p}{d t} = b v N. \quad (3.13)$$

Equations (5.32) and (3.13) provide the basis for the analysis of the plastic deformation dynamics.

Differentiating the applied stress \( \tau \) and the mobile dislocation density \( N \) with respect to the total resolved shear strain \( \varepsilon \) and writing the total derivative as

$$\frac{d ( \_ )}{d \varepsilon} = \frac{1}{\dot{\varepsilon}} \frac{d ( \_ )}{d t}, \quad (5.33)$$

gives

$$\frac{d \tau_a}{d \varepsilon} = \eta (1 - \frac{b v N}{\dot{\varepsilon}}), \quad (5.34a)$$

$$\frac{d N}{d \varepsilon} = \frac{1}{\dot{\varepsilon}} K_1 \tau_{eff} v N, \quad (5.34b)$$

$$K_1 = K \frac{8 \pi (1 - \nu)}{G b}, \quad (5.34c)$$

according to eqs. (5.32), (3.13) and (5.29), respectively. Integrating eqs. (5.34a) and (5.34b) gives the stress-strain curve of the material and the dependence of the dislocation
density on the total resolved shear strain. The elastic constant $\eta$ is calculated by fitting the initial slope of the experimental stress-strain curves.

Equations (5.34a) and (5.34b) are initially integrated to rederive the results of Suezawa et al. (1979) for high purity silicon and to investigate the validity of the model of Alexander and Haasen (1968). The initial value of the applied stress is zero and the initial dislocation density value is set for each experiment. The velocity equation (3.14) is rewritten as

$$v = U_0 \left( \frac{\tau_{\text{eff}}}{\tau_0} \right)^m \exp \left( -\frac{Q}{kT} \right),$$  \hspace{1cm} (5.35)

where $U_0 = 4.3 \times 10^4$ m/s, $\tau_0 = 10^7$ Pa, $m = 1.1$ and $Q = 2.17$ eV (Suezawa et al., 1979). Equations (5.11), (5.12) and (5.15) are used for the effective stress. This model is quite successful for the initial stages of deformation up to the lower yield region and gives accurate predictions of the upper yield point. Although the model predicts a lower yield point, the predictions are not accurate because the model does not include sufficient work hardening at high deformations.

To solve the above problems Suezawa et al. (1979) modified the expression for the effective stress based on Sumino's hypothesis for the steady state of deformation (Sumino, 1974) to

$$\tau_{\text{eff}} = \tau - A \sqrt{N} - C(\epsilon - \epsilon_{uy})^{1/2},$$  \hspace{1cm} (5.36)

where $\epsilon_{uy}$ is the strain at the upper yield point and $C$ depends on strain rate and temperature. Although there is a thermodynamic foundation for introducing the strain hardening function in the expression for the effective stress (Sumino, 1974; Suezawa et al., 1979), the corresponding micromechanical picture in the dislocation network is not clear. According to Sumino (1974) and Suezawa et al. (1979), the strain hardening function models the interactions between mobile and immobile dislocations. The form of the strain hardening function was attributed to a possible proportionality between the density of immobile dislocations and strain (Suezawa et al., 1979). Introduction of strain hardening functions in the definition of the effective stress is not adopted in the analysis presented. The stress-strain curves obtained from the integration of eqs. (5.34) with and without introducing work hardening are shown in Fig. 5.16.
Figure 5.16. Comparison between the predicted stress-strain curves in high purity silicon at $T = 800^\circ \text{C}$, $\dot{\varepsilon} = 1.2 \times 10^{-4} \text{ s}^{-1}$ and $N_0 = 10^4 \text{ cm}^{-2}$, with and without introducing a strain hardening function in the expression for the effective stress, eq. (5.36).
The velocity-stress relation derived in Section 5.1 and the constitutive models for the effective stress and dislocation multiplication derived in Section 5.2 are introduced for the simulation of dynamic deformation in oxygen-hardened silicon. Equations (5.34) are integrated numerically using the explicit Euler algorithm (Dahlquist and Björck, 1974) and a step size $\delta \varepsilon = 10^{-6}$. The integration is carried out at $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$, $T = 1073$ K and an initial dislocation density $N_0 = 10^6$ cm$^{-2}$ to simulate the experiments of Yonenaga et al. (1984).

The computed stress-strain curves are shown in Fig. 5.17 and are compared directly with the experimental data. The model is successful for the initial stages of deformation, from the elastic region to the upper yield region; the upper yield points for samples with five different oxygen concentrations are predicted within 10\% accuracy with respect to the experimental values. The upper yield stress provides a measure for the mechanical strength of the material. The results of Fig. 5.17 show the pronounced increase in the mechanical strength with respect to the high-purity material as the concentration of dissolved interstitial oxygen increases. This behavior explains the resistance of Czochralski silicon wafers to warpage during thermal cycling with respect to floating zone silicon wafers (Sumino, 1987a).

The dependence of the dislocation density on strain for the same samples and deformation conditions is presented in Fig. 5.18. The inflection points in the dependence of the dislocation density on strain correspond to the upper yield points in the stress-strain curves, because the condition $d\tau/d\varepsilon = 0$ at the upper yield point is equivalent to the condition $d^2N/d\varepsilon^2 = 0$. The $N = N(\varepsilon)$ curves also are used to calculate the critical resolved shear stress (CRSS) of the material; an analysis of the calculation of the CRSS is given in Section 5.3.2.

The dependence of the upper yield stress on the initial dislocation density for the same samples and deformation conditions is shown in Fig. 5.19. The results of the numerical calculations are fit using the relation

$$\tau_{uy} = C_{uy}^{1/3} \left[ \ln \left( \frac{C_0}{N_0} \right) \right]^{1/3},$$  \hspace{1cm} (5.37)

according to the theory of the upper yield point of Alexander and Haasen (1968). The
Figure 5.17. Comparison between the (a) predictions of the model, and (b) experimental data for the stress-strain curves at $T = 800^\circ$C, $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$ and $N_0 = 10^8$ cm$^{-2}$ for several oxygen concentrations.
Figure 5.18. Dependence of the mobile dislocation density on strain in crystals for several oxygen concentrations during deformation at $T = 800^\circ$C, $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$ and $N_0 = 10^6$ cm$^{-2}$.
Figure 5.19. Dependence of the upper yield stress on the initial dislocation density $N_0$ with varying oxygen concentrations during deformation at $T = 800^\circ$C and $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$. The dashed curves represent two-point fits according to eq. (5.37).
predictions of the model for the dependence of the upper yield stress \( \tau_{uy} \) on the oxygen concentration \( c_o \) are shown in Fig. 5.20 for \( T = 1073 \) K, \( N_u = 2 \times 10^5 \) and \( N_0 = 10^6 \) cm\(^{-2}\). The predictions of the model are in good agreement with the experimental data of Yonenaga et al. (1984). The dependence of \( \tau_{uy} \) is stronger in the high oxygen concentration range than in the low oxygen concentration range, in agreement with the experiment. This effect is interpreted to be a result of the exponential dependence of the dislocation multiplication rate constant \( K \) on the unlocking stress \( \tau_u \) given by eq. (5.30).

An exponential dependence of \( \tau_{uy} \) on \( c_o \) is suggested by the proportionality of \( \tau_u \) on \( c_o \) given by eq. (5.10), and the dependence of \( \tau_{uy} \) on \( K \), given by the simplified formula of Alexander and Haasen (1968), eq. (5.37), where \( C_{uy} \) is inversely proportional to \( K \).

The prediction of a lower yield point is demonstrated by the stress-strain curve shown as Fig. 5.21a for \( T = 1173 \) K, \( \dot{\varepsilon} = 1.1 \times 10^{-4} \) s\(^{-1} \), \( c_o = 1.5 \times 10^{17} \) cm\(^{-3} \) and \( N_0 = 10^6 \) cm\(^{-2} \). The variation of the predicted dislocation density \( N \) on strain \( \varepsilon \) for the same set of parameters is shown in Fig. 5.21b. The work hardening introduced by the stress, eq. (5.25), is not enough to successfully predict deformation at high strains because the complexity of the interactions in the dislocation network is not captured by the regular periodic array of \( 60^\circ \) dislocations used in our formulation. There are two reasons for this failure of the model at high strains. First, only interactions between parallel dislocations are taken into account by the model, while interactions between groups of dislocations, e.g. dislocation dipoles, that are formed at high dislocation densities are not included. Second, the equations used for the interaction between dislocations, e.g. eq. (5.16), are rigorously valid in the dilute limit, i.e. for low dislocation densities. For high dislocation densities, which is the case at high strains, corrections for the interaction between dislocations may become dominant. For strains higher than approximately five percent, the dependence of \( N \) on \( \varepsilon \) predicted by our model is successfully represented by a parabola. This result is analyzed further in Section 5.3.2. The predicted lower yield stress \( \tau_{ly} \) is proportional to the oxygen concentrations \( c_o \), as is demonstrated in Fig. 5.22 for \( T = 1173 \) K, \( \dot{\varepsilon} = 1.1 \times 10^{-4} \) s\(^{-1} \), and \( N_0 = 10^6 \) cm\(^{-2} \).
Figure 5.20. Comparison between the predictions of the model and experimental data of the dependence of the upper yield stress on the oxygen concentration during deformation at $T = 800^\circ$C and $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$ for two initial dislocation densities.
Figure 5.21. Dependence of (a) stress and (b) dislocation density on strain including stages of deformation past the yield region. The concentration of oxygen in the crystal is \( c_o = 1.5 \times 10^{17} \text{ cm}^{-3} \) and the deformation is carried out at \( T = 900^\circ \text{C} \) and \( \dot{\varepsilon} = 1.1 \times 10^{-4} \text{ s}^{-1} \). The dashed curve in (b) is a two-point parabolic fit.
Figure 5.22. Lower yield stress as a function of oxygen concentration during deformation at $T = 900^\circ\text{C}$ and $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$.
5.3.2. Estimation of the Critical Resolved Shear Stress

For the study of the deformation behavior of crystals with high initial dislocation densities, or generally with nonzero dislocation densities, we give an alternative definition for the CRSS compared to the conventional definitions presented in Section 2.1.2. This definition is motivated by the kinetics of dislocation multiplication and the growth of the dislocation population at low strains in the initial stage of deformation as is shown in Fig. 5.18. This growth is initially exponential, as is implied by eq. (5.34b). Exponential growth is approximately valid before the inflection point of the curve \( N = N(\varepsilon) \), which corresponds to the upper yield point, while past the inflection point the deviation from exponential growth is significant because of the increasing plastic strain relative to the elastic contribution. We define the CRSS as the stress in the region of the stress-strain curve below the upper yield point that corresponds to a critical strain defined as

\[
\varepsilon_{\text{CRSS}} \equiv \dot{\varepsilon} t_D,
\]

where \( t_D \) is the characteristic time for the initial exponential growth of the dislocation population, or the time required to increase the dislocation density from the initial value \( N_0 \) to \( N = 2N_0 \), i.e. the dislocation density doubling time.

The CRSS is calculated according to eq. (5.38) for the five silicon samples shown in Figs. 5.17 and 5.18. The CRSS values for these samples are 7.44, 15.55, 23.15, 32.68, and 42.61 MPa, respectively. These values are compared with the corresponding values for the upper yield stress which are 8.11, 17.34, 24.61, 33.38, and 42.65 MPa, respectively, and are always lower than the upper yield stresses. However, as the oxygen concentration increases and improves the mechanical strength of the material, the CRSS moves closer to the upper yield stress.

The yield stresses and the CRSS depend on the deformation conditions \((\dot{\varepsilon}, T)\) and the composition of the sample \((c_o)\). Here we give an approximate derivation of the CRSS in closed form. We approximate the effective stress for small deformations by the applied stress \(\tau\), assuming that the total internal stress \(\tau_i\) is negligible compared to \(\tau\). This assumption follows the approach of Section 3.3 in neglecting the back stress in the asymptotic
analysis. The validity of the approximation is based on the low value of the back stress \( A\sqrt{N} \) for low dislocation densities. Then, eq. (5.34b) becomes

\[
\frac{dN}{d\varepsilon} = \frac{1}{\dot{\varepsilon}} K_1 N \nu \tau.
\] (5.39)

As was mentioned in Section II, the dislocation velocity \( v \) is proportional to \( \tau \) for stresses higher than the unlocking stress \( \tau_u \); therefore,

\[ v = B_1(T)\tau. \] (5.40)

The temperature function \( B_1(T) \) is derived from the results for the dependence of the dislocation velocity on the applied stress. For example, in the high purity material eq. (5.3) implies that

\[ B_1(T) = B_0 \exp\left(-\frac{Q}{kT}\right). \] (5.40a)

In addition, neglecting the plastic contribution to the applied stress for small deformations gives

\[ \tau = \eta \varepsilon, \] (5.41)

according to eqs. (5.32) and (5.34a). Combining eqs. (5.40) and (5.41) and integrating eq. (5.39) gives the closed-form expression for the exponential growth of the dislocation population as

\[ N = N_0 \exp\left[\left(\frac{K_1 B_1 \eta^2}{3\dot{\varepsilon}}\right)\varepsilon^3\right]. \] (5.42)

Equation (5.42) implies that the critical strain \( \varepsilon_{CRSS} \), defined by eq. (5.38), is

\[ \varepsilon_{CRSS} = \left[\left(\frac{3\dot{\varepsilon}}{K_1 B_1 \eta^2}\right)\ln 2\right]^{1/3}. \] (5.43)

The dislocation density doubling time is given by eq. (5.38) as

\[ t_D \equiv \frac{\varepsilon_{CRSS}}{\dot{\varepsilon}} = \frac{1}{\dot{\varepsilon}} \left[\left(\frac{3\dot{\varepsilon}}{K_1 B_1 \eta^2}\right)\ln 2\right]^{1/3}. \] (5.44)
Combining eqs. (5.41) and (5.43) gives the CRSS as

$$\text{CRSS} = \eta \left[ \left( \frac{3 \dot{\varepsilon}}{K_1 B_1 \eta^2} \right) \ln 2 \right]^{1/3}. \quad (5.45)$$

Equation (5.45) is an approximate expression for the CRSS. The least satisfactory assumption used for the derivation of eq. (5.45) is the negligible plastic contribution to the total strain introduced by eq. (5.41). However, eq. (5.45) is satisfactory as an order-of-magnitude estimate.

A correlation for the dependence of the elastic constant $\eta$ on strain rate and temperature has been derived by fitting the results of Yonenaga and Sumino (1978) for high purity silicon. The correlation is accurate in the temperature range $1073 \leq T \leq 1123$ K and is given by

$$\eta = \alpha_0(\dot{\varepsilon}) \ G(T) \ \exp\left( \frac{Q_\alpha}{kT} \right), \quad (5.46)$$

where $Q_\alpha = 1.216$ eV and $\alpha_0 = 4.5632 \times 10^{-8}$ at $\dot{\varepsilon} = 1.2 \times 10^{-4}$ s$^{-1}$. Using eqs. (5.40a), (5.45) and (5.46) gives the dependence of CRSS on strain rate and temperature as

$$\text{CRSS} = \left( \frac{3 \ln 2}{K_1 B_0} \right)^{1/3} \left[ \alpha_0(\dot{\varepsilon}) \dot{\varepsilon} \right]^{1/3} G^{1/3}(T) \exp\left( \frac{Q + Q_\alpha}{3kT} \right), \quad (5.47)$$

where the temperature dependence of the shear modulus $G$ is given by eq. (5.17). At $T = 1073$ K and $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$ eq. (5.47) gives CRSS=12.66 MPa, which is 70% higher than the numerical value of 7.44 MPa for high purity silicon. However, eq. (5.47) describes the importance of rate processes and thermally activated processes for the determination of the CRSS.

The simplest approach for incorporating the effect of oxygen on the CRSS in silicon is to replace the multiplication rate constant $K_1$ by the constitutive expression, eq. (5.30). The approximation is that for applied stresses significantly higher than the unlocking stress $\tau_u$ the dependence of the dislocation velocity on the applied stress is the same as in the high purity material. This approximation is accurate only within a factor of two, according to the results presented in Section 5.1.4. Within these approximations, the CRSS in the
solute-hardened material is

\[
\text{CRSS} = \left( \frac{3 \ln 2}{K_{10} B_0} \right)^{1/3} \left[ \alpha_0 (\dot{\varepsilon}, c_o) \dot{\varepsilon} \right]^{1/3} G^{1/3} (T, c_o) \exp \left( \frac{Q + Q_o + \tau_u b^2 l}{3kT} \right),
\]

(5.48)

where \( K_{10} \equiv K_0 \cdot 8\pi (1 - \nu)/Gb \), the dependence of \( \eta \) on \( c_o \) is introduced in the factor \( \alpha_0 \), the dependence of \( G \) on \( c_o \) is given by eq. (5.19) and the dependence of \( \tau_u \) on \( c_o \) and \( T \) is given by eq. (5.10). The dominant dependence of CRSS on the oxygen concentration is introduced through the proportionality of \( \tau_u \) on \( c_o \). Therefore, eq. (5.48) implies an almost exponential increase of CRSS as a function of \( c_o \) at constant \( \dot{\varepsilon} \) and \( T \) with respect to the CRSS of the high purity material. This is consistent with the dependence of \( \tau_{uy} \) on \( c_o \) presented in the calculations in Section 5.3.1.

Although the analysis for the CRSS has been limited to silicon and oxygen-hardened silicon, the approximate expression (5.45) can be extended to all diamond/zincblende structure semiconductors by substituting \( m + 2 \) for the numerical factor (3) where \( m \) is the stress exponent in equation (2.5). Therefore, equation (5.45) gives the dependence of the CRSS on the multiplication rate constant as

\[
\text{CRSS} = \gamma_1 \cdot K_1^{-1/(m+2)},
\]

(5.49)

where \( \gamma_1 \) is a function of \( B_1, \eta, \) and \( m \) for the same deformation conditions. The exponent \( m \) is equal to 1.4 for InP (Müller et al., 1985) and 1.7 for GaAs (Yonenaga et al., 1987).

Expression (5.49) can be used to estimate the value of the multiplication rate constant of GaAs as it was mentioned in Section 3.2.2 based on the value of the multiplication rate constant of InP. According to eq. (5.49) the multiplication rate constant of GaAs is given by

\[
K_{1, \text{GaAs}} = K_{1, \text{InP}} \left( \frac{\text{CRSS}_{\text{InP}} / \gamma_{1, \text{InP}}}{\text{CRSS}_{\text{GaAs}} / \gamma_{1, \text{GaAs}}} \right)^{3.4/3.7}.
\]

(5.50)

The approximation is introduced that the values of \( \gamma_1 \) are similar for GaAs and InP. Using equations (5.40a) and (5.46) and the parameters of Table 3.1 estimates the accuracy of the approximation to a factor of two to three for the same deformation conditions for both materials at temperatures around 1000 K. Estimating \( \text{CRSS}_{\text{GaAs}} \approx 2.55 \) MPa.
(Guruswamy, 1987) and \( CRSS_{\text{InP}} \approx 1.58 \text{ MPa} \) (Müller et al., 1985) at the strain rate \( \dot{\varepsilon} = 10^{-4} \text{ s}^{-1} \), \( T = 973 \text{ K} \) and in the [123] crystal orientation gives \( K_{\text{GaAs}} \approx 0.2K_{\text{InP}} \).

5.3.3. Limiting Cases of the Constitutive Model of Plastic Deformation Dynamics

The models for the effective stress and dislocation multiplication developed in Section 5.2 give quantitatively meaningful predictions for the mechanical behavior of silicon when the dislocation density remains dilute, i.e. the direct interaction of dislocations is unimportant. This theory can be exploited to produce simpler expressions for \( \tau \) and \( N \) in two distinguished limits: the cases of large and small deformations relative to the strain at the upper yield point. These limits are described below.

5.3.3.1. Large Deformations

As the dynamic deformation experiments proceed to higher values of strain the density of mobile dislocations in the crystal increases. Excluding any strain hardening function from the model for the effective stress prevents the dislocation density from reaching a maximum steady-state value. Hence, the back stress \( A\sqrt{N} \) due to the dislocation interactions increases and becomes the dominant term in the expression for the internal stress, eq. (5.25). Alternatively, the unlocking stress does not depend on the dislocation density and becomes unimportant past the initial stage of deformation. Also, as the dislocation density grows the contribution to the internal stress due to the oxygen cloud-dislocation interaction is very small compared to the direct interaction of dislocations and becomes unimportant relative to the back stress. In the limit of high strains, this leads to a simplification of the deformation equations (5.34) to

\[
\frac{d\tau}{dt} \approx C_1 - C_2(\tau - A\sqrt{N})N, \tag{5.51a}
\]

\[
\frac{dN}{dt} \approx C_3(\tau - A\sqrt{N})^2 N, \tag{5.51b}
\]

where \( C_1 = \eta \dot{\varepsilon}, \ C_2 = \eta \dot{\varepsilon}b\nu/\tau_{\text{eff}} \) and \( C_3 = K_1\nu/\tau_{\text{eff}} \); these are constants under constant
strain rate and temperature, because the dislocation velocity is simply proportional to stress for steady dislocation motion according to the analysis of Section 5.1.4.

Equations (5.51) do not predict steady state values of dislocation density and stress, because there is no relation between $\tau$ and $N$ which sets both time derivatives to zero. This is not the case for static deformations, where an equilibrium condition exists. This is the condition of zero effective stress and it has been used as a quasi-steady state approximation for the plastic deformation dynamics of as-grown crystals in Section 3.3. However, the stress-strain curve of Fig. 5.21a suggests a quasi-steady state in stress for large deformations. In this case, setting the right side of eq. (5.51a) equal to zero gives

$$\tau = A\sqrt{N} + \frac{C_1}{C_2^2} \frac{1}{N}. \quad (5.52)$$

Substituting eq. (5.52) into eq. (5.51b) gives

$$\frac{dN}{dt} \approx \frac{C_3 C_1^2}{C_2^2} \frac{1}{N}, \quad (5.53)$$

and integrating with $\epsilon \equiv \dot{\epsilon} t$ gives

$$N^2 = N_0^2 + \frac{2C_3 C_1^2}{C_2^2} \frac{\epsilon}{\dot{\epsilon}}. \quad (5.54)$$

This parabolic relation between $N$ and $\epsilon$ has been used for fitting the results in Fig. 5.21b.

5.3.3.2. Small Deformations

A second limit of the constitutive model is more interesting, because it describes a limiting case of small deformation, where the model successfully gives quantitative predictions. The decrease of the dislocation multiplication rate constant $K$ as the oxygen concentration increases, predicted by eqs. (5.10) and (5.30) suggests the study of the asymptotic limit $K/K_0 \ll 1$. As $K$ goes to zero, the constant $C_3$ goes to zero and eqs. (5.51) become

$$N \simeq N_0, \quad (5.55a)$$
\[
\frac{d\tau_a}{dt} \simeq (C_1 + AC_2N_0^{3/2}) - (C_2N_0)\tau_a. \tag{5.55b}
\]

Integrating eq. (5.55b) gives an exponential dependence of stress on strain as

\[
\tau_a = \left(\frac{C_1 + AC_2N_0^{3/2}}{C_2N_0}\right)[1 - \exp\left(-\frac{C_2N_0}{\dot{\epsilon}}\right)], \tag{5.56}
\]

which leads to a steady state after a characteristic relaxation time

\[
t_{\text{rel}} \equiv \frac{1}{C_2N_0}. \tag{5.57}
\]

The limiting case of zero dislocation multiplication expressed by eqs. (5.56) and (5.57) corresponds to the well-known phenomenological model of a Maxwell viscoelastic solid (Krausz and Eyring, 1975). The Maxwell solid is modeled as a Hookean spring connected in series with a dashpot filled with a viscous liquid and gives the constitutive model

\[
\frac{d\tau}{dt} = C_{\text{el}} \frac{d\varepsilon}{dt} - \frac{\tau}{t_{\text{rel}}}, \tag{5.58}
\]

where \(C_{\text{el}}\) is the elastic constant of the solid. Integration of eq. (5.58) under conditions of constant strain rate, \(d\varepsilon/dt = \dot{\varepsilon}\), gives

\[
\tau = C_{\text{el}}\dot{\varepsilon}t_{\text{rel}}\left[1 - \exp\left(-\frac{\varepsilon}{\dot{\varepsilon}t_{\text{rel}}}\right)\right], \tag{5.59}
\]

which is the same form as eq. (5.56) with \(C_{\text{el}} \equiv \eta[1 + AN_0^{3/2}b(v/\tau_{\text{eff}})]\) and \(\eta\) the elastic constant in eqs. (5.34). The stress-strain curve for a silicon crystal with oxygen concentration \(c_0 = 1.5 \times 10^{17} \text{ cm}^{-3}\) at \(T = 1073 \text{ K}\), \(\dot{\varepsilon} = 1.1 \times 10^{-4} \text{ s}^{-1}\) and \(N_0 = 10^6 \text{ cm}^{-2}\) is shown in Fig. 5.23 together with the corresponding curve in the limit of the Maxwell solid \((K \to 0)\) with relaxation time \(t_{\text{rel}} = 149 \text{ s}\).

\section*{5.4. Summary and Discussion}

The study of the effects of oxygen on the mechanical strength of silicon crystals has been based on a self-consistent analysis. The analysis includes the effects of oxygen on the
Figure 5.23. Predictions of the deformation behavior between silicon with oxygen concentration $c_o = 1.5 \times 10^{17}$ cm$^{-3}$ from the model and of a Maxwell solid with $t_{rel} = 149$ s; $T = 800^\circ$C and $\dot{\varepsilon} = 1.1 \times 10^{-4}$ s$^{-1}$.
motion and locking of $60^\circ$ dislocations due to the elastic interaction between the dislocation and oxygen atoms in interstitial lattice sites and the formation of oxygen aggregates inside the dislocation core. Moreover, the classical models of Alexander and Haasen (1968) for the effective stress because of the internal stresses that are developed by the interactions between defects and of Johnston and Gilman (1960) for the rate of dislocation multiplication have been appropriately extended to include the effects of solute atoms in the crystal. The extension of this analysis to include electrically inactive impurities in other diamond/zincblende structure crystals is straightforward, as long as transport coefficients are known from experimental work or can be calculated by simulation and when models for the state of the impurities inside the dislocation core can be constructed. The analysis of the effects of electrically active impurities, i.e. donors and acceptors, is more difficult because the electrical interactions are at least of the same order-of-magnitude as the elastic interactions (Bullough and Newman, 1970; Nabarro, 1987). The magnitude of these interactions is crucial not only for the correct calculation of the dependence of dislocation velocity on applied stress at high stresses but also for the calculation of the unlocking stress and internal stresses.

The satisfactory prediction of the initial stage of the deformation of silicon requires the accurate calculation of the dislocation velocity and the unlocking stress at several temperatures. The unlocking stress appears both as a contribution to the total internal stress and to the activation energy for dislocation multiplication eq. (5.30). The expression of the unlocking stress used in the analysis is given by eq. (5.10) and the activation energy, $Q_c = 0.215$ eV, was derived from numerical calculations using the microscopic model of dislocation locking caused by oxygen. However, the temperature dependence of the unlocking stress can be explained by a simplified calculation. To do this the unlocking stress is modeled qualitatively as the stress required for a dislocation to reach a critical velocity that is equal to the drift velocity of the impurity. Using Einstein’s equation for the relation between the drift velocity and the isotropic diffusion coefficient and an activation energy for oxygen migration equal to the difference between the activation energy in the dislocation free material, $Q^o = 2.5$ eV (Needels et al., 1991), and the interaction energy $E_i = 0.5$ eV, enhanced migration of oxygen is predicted next to the core with an activation
energy of \( Q_m \approx Q^o - E_i = 2.0 \) eV. Subtracting \( Q_m \) from the activation energy for dislocation glide, \( Q = 2.2 \) eV, gives a value of \( Q_c \) equal to 0.2 eV, in good agreement with the value 0.215 eV of equation (5.10). However, the validity of eq. (5.10) is questionable for temperatures higher than about 1173 K because the one-dimensional oxygen clusters are not stable at higher temperatures (Needels et al., 1991).

The significant increase in the strength of oxygen-hardened silicon with respect to the high purity material is demonstrated in Fig. 5.17 and is caused mainly by the decreases of the dislocation velocity and of the multiplication rate constant with increasing oxygen concentration. The effect of oxygen on both the dislocation velocity and the internal stresses of eq. (5.25) is not significant for high applied stresses. The unlocking stress is proportional to \( c_o \) and reduces the effective stress significantly for low applied stresses. The increasing oxygen concentration also causes the multiplication rate constant \( K \) to decrease by as much as two orders-of-magnitude, depending on \( c_o \), and is responsible for the improvement of the mechanical properties of Czochralski silicon.

The constitutive models for the effective stress and dislocation multiplication rate, introduced in Section 5.2, fail to correctly predict work hardening for deformation past the lower yield region. However, the model is quite successful for small deformations, as shown in Figs. 5.17 and 5.20. The quantitative predictions of the upper yield points for various oxygen concentrations motivate the application of the model to several problems associated with the mechanical properties of Czochralski silicon, during both crystal growth and device processing. The stress and strain fields associated with such problems are within the initial stages of deformation. Nevertheless, modeling the effects of oxygen on the complex interactions in the dislocation network under large deformations may be very helpful to understanding the onset of polygonization in the Czochralski growth of very large diameter silicon crystals (Lin and Benson, 1987).

The dynamics of interstitial oxygen in silicon is studied as a case of point defect dynamics focusing on the effects of the migration of an electrically inactive impurity on dislocation dynamics. Quantitative modeling of impurity and dislocation dynamics is based on systematic analysis which is carried out over very different length scales, where the link between microscopic and macroscopic studies is provided by constitutive modeling. The
study of point defect dynamics is completed in Chapter 6 where the dynamics of intrinsic point defects in silicon is analyzed in the atomic scale. Systematic analysis of intrinsic point defect dynamics is considered as the first necessary step of an integrated methodology for modeling dislocation formation in single crystals in analogy with the analysis of impurity dynamics as a first step in the modeling of dislocation propagation and multiplication in single crystals with high impurity concentrations.
6. DYNAMICS OF INTRINSIC POINT DEFECTS IN SILICON

The objective of this chapter is to investigate the dynamics of intrinsic point defects in silicon by atomistic scale computer simulation and to examine the contributions of point defects to self-difusion and substitutional dopant diffusion. A systematic analysis is presented that is based on the calculation of structural, thermodynamic, and transport properties of the intrinsic point defects, i.e. vacancies and self-interstitials. The results of the simulations are compared with results from first-principle calculations and are used for the analysis and interpretation of experimental data. Finally, a methodology is developed for the atomistic calculation of intrinsic point defect properties that provides a robust computational framework for atomistic studies in crystals with larger defects, such as clusters of vacancies and self-interstitials.

A review of the current knowledge of the structure and properties of vacancies and self-interstitials in silicon is presented in Section 6.1. Point defect configurations and migration paths are defined and previous experimental data and theoretical results for intrinsic point defect properties and contributions to substitutional mass transport are presented. An integrated approach for the calculation of structural, thermodynamic, and transport properties of point defects in silicon based on atomistic simulations using the Stillinger-Weber interatomic potential (Stillinger and Weber, 1985) is developed in Section 6.2. The computational method combines lattice statics and Monte Carlo simulations with the quasi-harmonic approximation and the cumulant analysis of the enthalpy distribution function and yields values for the properties from zero temperature to the melting point. The results of our calculations for the structural, thermodynamic and transport properties of vacancies and self-interstitials including quantitative predictions for the thermal equilibrium concentrations and the diffusion coefficients of these defects are presented in Section 6.3. The results are summarized in Section 6.4 and the roles of vacancies and self-interstitials in silicon crystals are discussed.
6.1. Vacancies and self-interstitials in silicon

Vacancies and self-interstitials are intrinsic point defects or native lattice defects in crystalline solids because of their point-like character and because they exist in nonzero concentrations in every pure crystal for temperatures above absolute zero (Flynn, 1972). Intrinsic point defects in crystals are tremendously important because these defects provide the vehicles for substitutional mass transport, i.e. self-diffusion and substitutional impurity diffusion (Flynn, 1972). In silicon crystals substitutional impurities such as Au, Pt, the Group-III p-type dopants B, Al, Ga and the Group-V n-type dopants P, As, Sb migrate following mechanisms that involve both vacancies and self-interstitials (Tan and Gösele, 1985; Fahey et al., 1989). These dopants and impurities play important roles in modifying the mechanical, optical, and electronic properties of the intrinsic semiconductor. The need to control the distribution of dopants and impurities in the crystal during device processing emphasizes the role of the intrinsic point defects (Fahey et al., 1989).

6.1.1. Structural Configurations and Migration Paths

The single vacancy (V) is defined simply as an empty lattice site. In silicon the vacancy exists in several charge states (V\(^-\), V\(^0\), V\(^+\), V\(^{2+}\)) depending on how the resultant unsatisfied covalent bonds are reconfigured to accommodate the vacancy defect in the lattice. A self-interstitial is an extra silicon atom that resides in one of the interstices of the silicon lattice. The self-interstitial also exists in several charge states (I\(^0\), I\(^+\), I\(^{2+}\)) depending on the electron density distribution in the vicinity of the extra atom.

There are several structural configurations of point defects. A split-vacancy (SV) is defined as the point defect configuration that places one of the nearest-neighbors of a vacant site at the midpoint of the distance between the vacant site and the original lattice position. The atom at the SV configuration has a six-fold coordination compared to the three-fold coordination of the nearest-neighbors of the vacancy (Lannoo and Bourgoin, 1982). The split-vacancy is the classical saddle point configuration during a vacancy jump (Lannoo and Bourgoin, 1982).

Among various interstitial configurations the most important are: the tetrahedral (I\(_T\)
at site T), the hexagonal (I_H at site H), the bond-centered (I_B at site B) and the [100] split or dumbbell (I_D at configuration D) interstitials. (Lanno and Bourgoin, 1982). The tetrahedral site is at a distance equal to one bond length from a host lattice site in the direction defined by the host lattice site and one of its nearest neighbors. I_T is a high-symmetry configuration with four-fold coordination, i.e. the coordination of host lattice atoms. The hexagonal site is the midpoint between two nearest-neighbor tetrahedral sites, i.e. the classical saddle point of tetrahedral interstitial migration; the I_H site has six-fold coordination. The bond-centered site is the midpoint between two nearest-neighbor lattice sites, i.e. it is located at the center of the bond length between two neighboring atoms and it has a two-fold coordination. The dumbbell interstitial consists of the extra atom at a distance equal to half a bond length from a lattice site with the atom that occupied the lattice site being pushed in the [100] direction at a distance equal to one bond length from the extra atom. These unrelaxed interstitial configurations are shown in Fig. 6.1, as projected on the (011̅) plane of the diamond-cubic lattice.

Self-diffusion and substitutional dopant diffusion that involve vacancies consist of atomic jumps to vacant sites. In the case of the interstitial we distinguish between two types of diffusion: channel diffusion and interstitialc diffusion (Flynn, 1972). Channel diffusion occurs via simple paths which do not involve exchange with atoms at lattice sites. Such paths are the TH path, where interstitial migration occurs through tetrahedral and hexagonal sites and the BD path where migration occurs through bond-centered and dumbbell sites. Interstitialc paths involve exchange with atoms at lattice sites providing a mechanism for self-diffusion. For example, a tetrahedral interstitial can move in the [100] direction toward one of its second neighbors and push that atom into the next channel while taking its place (Car et al., 1984). Another interstitialc path (TB) has a tetrahedral interstitial moving in the [111] direction toward a nearest neighbor, pushing on two atoms and eventually ejecting one of them into the next channel (Car et al., 1984). All the simple and interstitialc migration paths are shown in Fig. 6.1.

The concept of an extended point defect also has been introduced in the literature (Seeger and Chik, 1968). An extended point defect is different from a localized lattice point defect in that it is not nearly as localized with respect to a specific site of the
Figure 6.1. Self-interstitial configurations in a (110) plane of a silicon lattice. The tetrahedral (T), hexagonal (H) and bond-centered (B) sites and the dumbbell configuration (D) are shown. The arrows denote simple migration paths which involve T and H sites (TH), and B and D sites (BD) and interstitialcy migration paths which involve T and B sites (TB), and T and D sites (TD).
unrelaxed lattice. The extended vacancy is defined as the defect involving the interactions of \( N - 1 \) silicon atoms about \( N \) lattice sites, and similarly the extended self-interstitial as the defect involving the interactions of \( N + 1 \) silicon atoms about \( N \) lattice sites. The lattice surrounding a point defect is distorted by the presence of the defect. An extended point defect involves changes in the lattice over many lattice sites. There is no definite transition from a point-like defect to an extended defect (Fahey et al., 1989). When the defect loses its point-like character it is not amenable to any kind of simple analysis. No work has been done in developing even an elementary model of migration of an extended defect.

6.1.2. Experimental Measurements of Intrinsic Point Defect Properties

Thermal equilibrium concentrations and diffusion coefficients of vacancies and self-interstitials are the point defect properties of immediate interest in the analysis of self-diffusion and substitutional dopant diffusion (Tan and Gösele, 1985; Fahey et al., 1989). No experiment has definitely measured the equilibrium concentrations \( c^{eq}_V \) and \( c^{eq}_I \) or the diffusion coefficients \( D_V \) and \( D_I \) of vacancies and self-interstitials, respectively. Therefore, accurate determination of these quantities is considered one of the most significant developments in understanding the roles of defects in silicon (Fahey et al., 1989). The absence of direct experimental measurements of equilibrium concentrations and diffusivities of intrinsic point defects has led to controversy over the role of intrinsic point defects in diffusion. Estimated values of the individual factors \((c^{eq}_V, D_V, c^{eq}_I, D_I)\) spread over many orders-of-magnitude at a given temperature (Fahey et al., 1989; Tan and Gösele, 1985).

Native monovacancies have been observed at low temperatures (Watkins et al., 1979) by electron paramagnetic resonance (EPR) and deep-level transient spectroscopy (DLTS) after electron irradiation of silicon. Self-interstitials have not been directly observed by EPR, but their presence has been inferred by reactions which occur during electron irradiation. The important question at high temperatures is whether interstitials or vacancies dominate diffusion processes. Most of the experimental work at high temperatures focuses on measurements of dopant concentration profiles (Fahey et al., 1989; Tan and Gösele,
1985). All the results that have been reported in these studies are based on indirect analyses of the diffusion data. The most valuable conclusion of these studies is that both vacancies and self-interstitials contribute to substitutional transport in silicon. The details of these analyses are reviewed separately in Section 6.1.3.

Estimates of the equilibrium atomic fractions and diffusion coefficients of vacancies and self-interstitials are derived based on the equations

$$c_{X}^{eq} = \theta_{X} \exp\left(\frac{S_{X}^{f}}{k}\right) \exp\left(-\frac{H_{X}^{f}}{kT}\right); \quad X = V, I$$  \hspace{1cm} (6.1)

$$D_{X} = f_{X} \exp\left(\frac{S_{X}^{m}}{k}\right) \exp\left(-\frac{H_{X}^{m}}{kT}\right); \quad X = V, I$$  \hspace{1cm} (6.2)

where \((H_{X}^{f}, S_{X}^{f}, H_{X}^{m}, S_{X}^{m})\) are the enthalpies and vibrational entropies of formation and migration, respectively, and \(k\) is Boltzmann’s constant. The factor \(\theta_{X}\) depends on the defect configuration and the constant \(f_{X}\) depends on the geometry of the lattice and the frequency of lattice vibrations. Many experimental techniques focus on the measurement of enthalpies and entropies of point defect formation and migration. Rigorous derivation of eqs. (6.1) and (6.2) based on thermodynamic analysis in the dilute limit and jump rate theory, respectively are presented in Section 6.2. These equations are generally valid for neutral point defects. Fahey et al. (1989) have reviewed the experimental determination of energy levels for the calculation of equilibrium concentrations of charge states.

The positron annihilation technique is the only method that is being applied for measurements of point defect formation properties at elevated temperatures and is considered the most direct experimental approach for determining equilibrium concentrations (Fahey et al., 1989). Positron lifetimes attributable to positrons annihilated in monovacancies are observed directly at \(T > 1450\) K (Dannefaer et al., 1986). The measured lifetimes are combined with a model for positron trapping by vacancies and the vacancy formation energy is measured as a slope of the trapping parameter as a function of inverse temperature (Dannefaer et al., 1986). The analysis of Dannefaer et al. (1986) gives a formation enthalpy of 3.6 eV for the neutral vacancy and a formation entropy between 6 k and 10 k. The analysis of this data has been criticized by Van Vechten (1987), who proposed that a
correct interpretation of the positron data yields \( H_V^f = 2.6 \) eV and \( S_V^f = 2.9 \) k. The origin of the disagreement between the two analyses is the possible association of the positron lifetime with divacancies or larger vacancy clusters, rather than monovacancies alone and the derivation of the magnitude of the trapping cross section (Van Vechten, 1986).

Migration rates of point defects in silicon have been measured by creating point defects in samples by irradiation with high-energy electrons, photons, or ions at temperatures low enough that the radiation-generated defects cannot migrate. Since the energy for creating the excess defects is provided externally, migration energies of intrinsic point defects are deduced by monitoring the annealing of the excess defect concentration at high temperatures (Fahey et al., 1989). In these irradiation experiments, vacancies and interstitials are produced in equal pairs. Electron irradiation experiments at cryogenic temperatures have shown that isolated vacancies are found frozen in silicon at \( T = 4.2 \) K, while interstitials are capable of athermal migration being mobile even at such low temperatures (Watkins, 1975). The experiments for the migration energies of the vacancies left behind in the radiation experiments derived after annealing at temperatures up to about 200 K are 0.33 eV, 0.45 eV and less than 0.18 eV for the \( V^0 \), \( V^{2+} \) and \( V^{2-} \) configurations, respectively (Watkins, 1975; Watkins et al., 1979; Newton et al., 1983). Experimental results for the migration energies of vacancies and interstitials created by ion implantation and proton bombardment have been reviewed by Fahey et al. (1989).

6.1.3. Self-Diffusion and Dopant Diffusion

The self-diffusion coefficient \( D_{SD} \) characterizes the transport of the host atoms in the crystal and is given in general by the Arrhenius expression

\[
D_{SD} = D_{SD}^0 \exp \left( -\frac{Q_{SD}}{kT} \right).
\] (6.3)

The experimentally observed values of the preexponential factor \( D_{SD}^0 \) for silicon are of order \( 10^3 \) cm\(^2\)/s (Fahey et al., 1989). These values are large compared to the corresponding values in metals and motivated the proposal of extended point defects in silicon (Seeger and Chik, 1968). Measurements of self-diffusion in silicon using radioactive tracers, Ge
as a tracer, and nuclear profiling techniques are reviewed by Fahey et al. (1989). The measured values of $Q_{SD}$ using these techniques at $T > 1200$ K are between 4 and 5 eV.

Assuming that both vacancy and interstitialcy diffusion mechanisms govern the migration of host silicon atoms, the self-diffusion coefficient is linearly decomposed into contributions from vacancies and self-interstitials (Tan and Gösele, 1985; Fahey et al., 1989) as

$$D_{SD} = c_{V}^{eq} D_{V} + c_{I}^{eq} D_{I}.$$ \hspace{1cm} (6.4)

The relative role of vacancies and self-interstitials in self-diffusion and substitutional dopant diffusion, i.e. the relative magnitude of the terms $c_{V}^{eq} D_{V}$ and $c_{I}^{eq} D_{I}$ has been a major controversy in the physics of point defects in silicon. The enhancement or retardation of dopant diffusion due to surface oxidation partially resolved this issue (Tan and Gösele, 1985). High temperature surface oxidation of silicon injects excess self-interstitials which lead to the nucleation and growth of interstitial-type stacking faults (Hu, 1991) and to enhanced diffusion of the substitutional dopants B, P and As, as well as retarded diffusion of Sb. A systematic analysis of stacking fault growth coupled with the effects on dopant diffusion led Hu (1974) to the conclusion that both vacancies and self-interstitials contribute to dopant diffusion in silicon.

Quantitative predictions for the contributions of vacancies and self-interstitials to self-diffusion have been based on the analysis of gold diffusion in silicon and experimental measurements of gold concentration profiles (Tan and Gösele, 1985). Diffusion of gold from the silicon surface into the bulk is governed by the one-dimensional transient equation of species conservation

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left( D_{eff} \frac{\partial c}{\partial x} \right),$$ \hspace{1cm} (6.5)

where the $x$-direction is perpendicular to the surface and $D_{eff}$ is the effective diffusivity of gold which, in general, depends on the gold concentration $c$. Gold in silicon is mainly dissolved substitutionally but moves by an interstitial mechanism (Tan and Gösele, 1985), resulting in fast diffusion rates at high temperatures. The move from substitutional sites ($Au_{s}$) to interstitial sites ($Au_{i}$) requires intrinsic point defects and is described by two
mechanisms written in reaction form as

\[ \text{Au}_i + V \rightleftharpoons \text{Au}_s, \]  
\[ \text{Au}_i \rightleftharpoons \text{Au}_s + I. \]

The first reaction describes the Frank-Turnbull mechanism (Frank and Turnbull, 1956) and the second reaction describes the kick-out mechanism (Gösele et al., 1980).

Introducing the intrinsic point defect contributions in the effective diffusivity (Tan and Gösele, 1985) gives

\[ D_{\text{eff}} = \frac{cV}{c} D_V + \frac{cI}{c} D_I. \]

Assuming local dynamical equilibrium between I, V, Au\(_i\) and Au\(_s\) in reactions (6.6) gives \( D_{\text{eff}} \) as a function of \( c_i^{eq} D_V \), \( c_i^{eq} D_I \), \( c \), and the solubility of gold in silicon \( c_s^{eq} \) (Tan and Gösele, 1985). Assuming that only the Frank-Turnbull mechanism (6.6a) operates gives an effective diffusivity independent of \( c \) and proportional to \( c_i^{eq} D_V \). In such a case \( D_{\text{eff}} \) is a constant and eq. (6.5) with the appropriate boundary conditions has a similarity solution expressed by a complementary error function of the similarity variable \( \xi = x/\sqrt{4D_{\text{eff}}t} \) (Bird et al., 1960). Fitting the experimentally measured gold profiles with the error function solution gives the value of \( c_i^{eq} D_V \).

Assuming that only the kick-out mechanism (6.6b) operates gives an effective diffusivity proportional to \( c_i^{eq} D_I c^{-2} \). Fitting the concentration profiles to the numerical solution of eq. (6.5) provides an estimate for \( c_i^{eq} D_I \) (Tan and Gösele, 1985). A complete analysis of gold diffusion in silicon is presented by Morehead (1988).

In spite of the uncertainties associated with the analyses of gold diffusion data (Tan and Gösele, 1985) all the resulting Arrhenius expressions have activation energies between 4 and 5 eV for both \( c_i^{eq} D_V \) and \( c_i^{eq} D_I \) and preexponential factors in of \( O(10^3) \) cm\(^2\)/s for \( c_i^{eq} D_I \) and of \( O(1) - O(10) \) cm\(^2\)/s for \( c_i^{eq} D_V \) (Tan and Gösele, 1985; Morehead, 1988). These results are in agreement with the direct measurements of self-diffusion and suggest that both vacancies and self-interstitials contribute to self-diffusion and substitutional dopant diffusion.
6.1.4. First-Principles Calculations of Intrinsic Point Defect Properties

First-principles calculations of intrinsic point defect properties provide the theoretical basis for a complete description of atomic diffusion processes in a crystal. In a first-principle calculation the Schrödinger equation (Ashcroft and Mermin, 1976) is solved self-consistently without any free parameters. Significant progress in theoretical developments for describing point defects in semiconductors has occured in the past decade (Pantelides, 1990). The Green’s function approach was initially used (Baraff and Schluter, 1978; Bernholc et al., 1978) to carry out first-principles calculations of the charge density and energy levels of point defects at the same level of sophistication and accuracy that was already possible for perfect bulk crystals and surfaces.

Modified Green’s function analyses (Car et al., 1984, 1985) and supercell calculations in a momentum space formulation (Bar-Yam and Joannopoulos, 1984a, b) for total energy (Ihm et al., 1979) followed providing a comprehensive picture of the nature and the energetics of intrinsic point defects in silicon. In these analyses norm-conserving pseudopotentials (Hamann et al., 1979) with plane-wave basis sets are used and electronic structure calculations are performed using density-functional theory within the local density approximation (Ceperley and Alder, 1980; Perdew and Zunger, 1981). Special \( k \) points in the irreducible part of the Brillouin zone are used for total energy integrations in the momentum space formalism (Chadi and Cohen, 1973).

Car et al. (1984, 1985) calculated the formation enthalpies of the vacancy and the self-interstitial for various charge states \( (V^{-}, V^{0}, V^{+}, V^{2+}, I^{0}, I^{+}, I^{2+}) \) as a function of the position of the Fermi level \( E_{F} \) in the band gap. The calculations showed that the equilibrium point defect configuration depends on the charge state and the position of the Fermi level and that vacancies and self-interstitials are negative-\( U \) centers, i.e. \( V^{+} \) and \( I^{+} \) are not the stable charge states for any value of \( E_{F} \). The calculations proved that simple configurations of the vacancy and the self-interstitial have comparable formation energies, both around 4 to 5 eV. Migration energies were small for both (Car et al., 1984; Pantelides et al., 1984), in agreement with experimental data. The self-interstitial is capable of athermal migration by the successive capture of electrons and holes which change the
charge state and, in turn, the stable site (Car et al., 1984, 1985; Pantelides et al., 1984; Bar-Yam and Joannopoulos, 1984a, b). Therefore, the total energy calculations concluded that both vacancies and self-interstitials have large formation energies and contribute to self-diffusion with an activation energy between 4 and 5 eV in agreement with the experimental observations (Seeger and Chik, 1968; Fahey et al., 1989).

The concerted exchange mechanism for diffusion of substitutional atoms was proposed by Pandey (1986) and is diffusion mechanism that does not involve vacancies and self-interstitials Total energy calculations (Pandey, 1986) proved that the energy barrier accompanying the exchange of two neighboring atoms is 4.3 eV. This suggests that the contribution of the concerted exchange mechanism to self-diffusion may be comparable with the contributions from the vacancy mechanism and the interstitialcy mechanism. However, experiments using barried layers of germanium in silicon suggest that injection of either vacancies or self-interstitials enhances the diffusion coefficient (Fahey et al., 1989). These results imply that both defects mediate self-diffusion and substitutional impurity diffusion, while the role of the concerted exchange mechanism is minimal since this mechanism does not depend on point defects. The different contributions to self-diffusion from mechanisms with similar energy barriers have been explained based on the differences in the corresponding preexponential factors (Pantelides, 1990).

Full dynamical first-principles calculations based on the Car-Parrinello method (Car and Parrinello, 1985) have shed new light in microscopic diffusion processes in semiconductors. This methodology which combines density-functional theory with molecular dynamics provides a unified simulation approach using supercells that contain up to $O(10^2)$ atoms and employing $O(10^2)$ time steps (Buda et al., 1989). The technique has been successfully applied to the calculation of hydrogen diffusivity in silicon (Buda et al., 1989; Chiarotti et al., 1990) and studies of surfaces (Needels et al., 1987) and impurities (Needels et al., 1991). The Car-Parrinello method is the most promising methodology for the in-depth investigation of microscopic dynamics in solids.
6.2. Integrated Atomistic Approach for Calculation of Intrinsic Point Defect Properties in Silicon

Atomistic simulations based on empirical or semi-empirical atomic models is a complementary approach to first-principles calculations in the investigation of point defect properties (Lannoo and Bourgoine, 1982; Batra et al., 1987; Baskes et al., 1989). Although empirical atomic models convey no information about the electronic structure, they provide useful tools for point defect studies. The use of large-size periodically repeated supercells containing clusters of hundreds or thousands of atoms are suitable to account for long-range lattice relaxations. Empirical atomic models, such as the Keating model (Keating, 1966) have been included in an otherwise first-principles approach to account for the long-range lattice relaxation, where the relaxation of the nearest neighbors of the point defect is included in a self-consistent way (Car et al., 1984). Furthermore, atomistic simulation is suitable for diffusion studies at finite temperatures under equilibrium conditions, where the necessary energy for the migration of the point defect is supplied by the lattice in the form of phonons.

The objective of this Section is to present the systematic development of a computational approach for the internally-consistent calculation of equilibrium and transport properties of intrinsic point defects in silicon based on atomistic simulation using empirical interatomic potentials. This computational framework will be used for structural and thermodynamic studies of larger defects, such as clusters of vacancies and self-interstitials, by providing the necessary parameters associated with the single point defects.

Application of atomistic computational analysis in the study of point defects in silicon is motivated by two objectives. The first is the investigation of the relaxed structure and the stability of atomic configurations that include intrinsic point defects and the mechanisms of point defect migration. The second is the calculation of formation and migration properties of intrinsic point defects both at the ground state and at finite temperatures. These calculations lead to the determination of the thermal equilibrium concentrations and the diffusion coefficients of vacancies and self-interstitials following the thermodynamic analysis for the single point defect distribution in the dilute limit (Girifalco, 1973;
Lannoo and Bourgoin, 1982) and the jump rate theory of point defect migration (Girifalco, 1973; Lannoo and Bourgoin, 1982), respectively. Therefore, the calculations provide all the input for constitutive modeling of self-diffusion and substitutional dopant diffusion in a direct and internally-consistent way. However, the quantitative predictions of this approach are limited by the range of validity of the empirical atomic model which emphasizes the importance of the use of a reliable interatomic potential.

The calculational method that we have developed is based on a combination of lattice statics and Monte Carlo simulations with lattice dynamics methods and analyses of distribution functions. An outline of these methods is given by the flow chart of Fig. 6.2. The details of the theories, models and computational procedures are given in Sections 6.2.1 - 6.2.6.

6.2.1. Atomic Model for Crystalline Silicon

Atomic models for monatomic systems are potential energy functions which describe interactions among \( N \) identical particles. In general, any potential energy function \( \Phi \) can be decomposed into one-body, two-body and many-body contributions (Allen and Tildesley, 1989) as

\[
\Phi(1, \ldots, N) = \sum_{i=1}^{N} v_1(i) + \sum_i \sum_{j>i} v_2(i, j) + \sum_i \sum_{j>i} \sum_{k>j>i} v_3(i, j, k) + \cdots + v_N(1, \ldots, N). \tag{6.8}
\]

The components \( v_n \) in the decomposition of eq. (6.8) converge quickly to zero with increasing \( n \). The single-particle potential \( v_1 \) normally describes interactions with walls or external forces to which the system is subject. In the absence of these forces the expansion begins with the pair-interaction terms.

Pairwise additive central force potentials have been traditionally used for decades as atomic models in condensed matter physics (Hansen and McDonald, 1986; Allen and Tildesley, 1989). The famous 6-12 Lennard-Jones potential (McQuarrie, 1976)

\[
v_2(i, j) = v_{LJ}(r_{ij}) = -4\epsilon \left( \frac{\sigma}{r_{ij}} \right)^{12} + \left( \frac{\sigma}{r_{ij}} \right)^6, \tag{6.9}
\]

where \( \epsilon \) and \( \sigma \) are the energy and length scales, and \( r_{ij} \) is the interatomic distance between particle \( i \) and particle \( j \), has been typically used for the study of liquified model gases such
Figure 6.2. Integrated approach for atomistic calculations of intrinsic point defect properties in silicon.
as Ar, Kr, and Xe. For solid state modeling the Morse potential

\[ v_2(i,j) = v_M(r_{ij}) = \epsilon \{ \exp[-2a(r_{ij} - r_0)] - 2 \exp[-a(r_{ij} - r_0)] \}, \quad (6.10) \]

where \( \epsilon \) is the energy scale and \( a \) and \( r_0 \) are length scales has been traditionally used for many metallic and covalent systems such as silicon (Lanno and Bourgoin, 1982).

Crystalline silicon consists of atoms held in place by strong and directional covalent bonds (Burns, 1985). For a quantitatively reliable description of this structure the corresponding potential energy \( \Phi \) is approximated by a combination of triplet and pair potentials, \( v_3 \) and \( v_2 \). Many three-body potentials have been proposed for silicon (Stillinger and Weber, 1985; Tersoff, 1986, 1988, 1989; Biswas and Hamann, 1985; Baskes, 1987), reflecting the needs for accurate and computationally efficient atomic modeling of silicon. The functional details of the interatomic potential of Stillinger and Weber (1985) are given here only, because this is the interatomic potential used in our calculations for perfect and imperfect crystalline silicon.

The Stillinger-Weber (SW) potential (Stillinger and Weber, 1985) combines the potential energy functions

\[
\begin{align*}
  v_2(r_{ij}) &= \epsilon f_2(r_{ij}) \\
  v_3(r_i, r_j, r_k) &= \epsilon f_3(r_i/\sigma, r_j/\sigma, r_k/\sigma),
\end{align*}
\]

(6.11a)
(6.11b)

where the energy scale \( \epsilon \) is chosen to give \( f_2 \) a depth of -1, the length scale \( \sigma \) is chosen to make \( f_2(2^{1/6}) \) vanish and \( r_i \) denotes the position vector of particle \( i \). The function \( f_3 \) possesses full translational and rotational symmetry. The functional form of the Keating potential (Keating, 1966) which gives a good description of elastic properties and phonon spectra is taken into account in the construction of the SW potential. However, the final form of the SW functions becomes much more involved in order to achieve a satisfactory description of the short-range order and of the atom-exchanging diffusive motion in the liquid phase (Stillinger and Weber, 1985).

The reduced SW pair potential is given by the five-parameter function

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

(6.12)
where \( A, B, p, \) and \( a \) are positive numbers. The function \( f_2 \) goes smoothly to zero without discontinuities in derivatives of any order. The three-body term is given by

\[
f_3(r_i, r_j, r_k) = h(r_{ij}, r_{ik}, \theta_{ijk}) + h(r_{ji}, r_{jk}, \theta_{ijk}) + h(r_{ki}, r_{kj}, \theta_{ijk}),
\]

(6.13a)

where \( h \) is the two-parameter function

\[
h(r_{ij}, r_{ik}, \theta_{ijk}) = \begin{cases} 
\lambda \exp \left[ \gamma (r_{ij} - a)^{-1} + \gamma (r_{ik} - a)^{-1} \right] (\cos \theta_{ijk} + \frac{1}{3})^2, & r_{ij}, r_{ik} < a; \\
0, & \text{otherwise}, 
\end{cases}
\]

(6.13b)

where \( \theta_{ijk} \) is the angle between \( r_j \) and \( r_k \) subtended at vertex \( i \), and \( a \) is the same smooth cutoff distance used for function \( f_2 \). Equation (6.13b) shows that the ideal tetrahedral angle \( \theta_t \) given by

\[
\cos \theta_t = -\frac{1}{3} \iff \theta_t = 109.47^\circ,
\]

(6.14)

makes the trigonometric portion of function \( h \) favor pairs of bonds emanating from vertex \( i \) that have the tetrahedral geometry.

Using eqs. (6.12) and (6.13) to fit thermophysical properties of the solid and the liquid state provides the most satisfactory values for the seven parameters \( (A, B, p, q, a, \lambda, \gamma) \). This parameter set is used in our simulations and is given in Table 6.1 together with the values of the length and energy scales. The reduced pair potential \( f_2 \) is plotted as a function of the interatomic distance in Fig. 6.3. Eqs. (6.12) and (6.13) with the parameter set of Table 6.1 guarantee that the diamond structure is indeed the most stable periodic arrangement of particles among other cubic structures. This is shown in Fig. 6.4 where the lattice energy

\[
\Phi/\epsilon = \sum_i \sum_{j>i} f_2(i,j) + \sum_i \sum_{j>i} \sum_{k>j>i} f_3(i,j,k)
\]

(6.15)

is plotted as function of density for the simple (sc), face-centered (fcc), body-centered (bcc), diamond (dc) cubic structures. The corresponding lattice sums of eq. (6.15) using \( N = 216, N = 256, N = 250, \) and \( N = 216 \) particles per periodic supercell, respectively. Obviously, the size of the system does not affect these calculations. The main reason
Figure 6.3. Reduced pair interaction of the interatomic potential of Stillinger and Weber.
Figure 6.4. Density dependence of the lattice energy for cubic lattice structures of silicon according to the Stillinger-Weber interatomic potential.
for using interatomic potentials that keep up to three-body terms in the decomposition of eq. (6.8) is the computational efficiency in our calculations. Moreover, other interatomic potentials with \( n \)-body terms (\( n \geq 3 \)) (Bolding and Andersen, 1990; Chelikowsky et al., 1989; Chelikowsky and Phillips, 1989) are better models for the study of clusters of a few atoms, but they have not been extensively used for calculation of bulk solid state and liquid state properties. The SW atomic model is preferred among other three-body potentials because it gives a very good prediction of the melting temperature \( T_m \) compared with the experimental value \( T_{m_n} = 1683 \text{ K} \) (Broughton and Li, 1987). The predictions of at least two independent studies (Broughton and Li, 1987; Phillpot et al., 1989) give \( T_m \sim 1690 \text{ K} \), an excellent simulation result. Accurate prediction of the melting temperature by an atomic model is a necessary significant feature for studies of imperfect crystalline phases at high temperatures. Furthermore, the SW potential was found to predict formation energies of relaxed strictly localized self-interstitial configurations in agreement with self-consistent field total energy calculations (Batra et al., 1987).
6.2.2. Atomistic Simulation Methods

Calculation of structural, thermodynamic and transport properties of perfect and imperfect crystalline silicon crystals are carried out using Monte Carlo and lattice statics simulation techniques. Lattice dynamics techniques also are used to calculate contributions to the free energy of the crystals. Monte Carlo simulations are carried out to calculate equilibrium ensemble averages, that correspond to thermodynamic properties of the system at finite temperatures by appropriate sampling of the configurational space. Lattice statics simulations based on energy minimization techniques are used to calculate relaxed atomic configurations that correspond to the minimum energy state (ground state) of the system at $T = 0$ K.

6.2.2.1. Metropolis Importance Sampling and Canonical Ensemble

Monte Carlo

The canonical $(NVT)$ ensemble is the statistical ensemble of systems with fixed number of particles $N$, volume $V$, and temperature $T$. The canonical ensemble average of a thermodynamic property is given in quasi-classical form (McQuarrie, 1976; Hansen and McDonald, 1986; Allen and Tildesley, 1989) by

$$
(A)_{NVT} = \frac{\int d\mathbf{r}^N \int d\mathbf{p}^N \exp(-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)/kT) A(\mathbf{r}^N, \mathbf{p}^N)}{\int d\mathbf{r}^N \int d\mathbf{p}^N \exp(-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)/kT)},
$$

(6.16)

where $\mathbf{r}^N$ and $\mathbf{p}^N$ denote the coordinates and the momenta of the $N$ particles of the system and $\mathcal{H}$ is the Hamiltonian of the system. For a monatomic system of $N$ particles the Hamiltonian is given by the sum of kinetic and potential energy as

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

(6.17)

The canonical ensemble partition function for an atomic system is written in quasi-classical form as

$$
Q_{NVT} = \frac{1}{N!} \frac{1}{\hbar^{3N}} \int d\mathbf{r}^N \int d\mathbf{p}^N \exp(-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)/kT),
$$

(6.18)
where \( h \) is Planck's constant. Eq. (6.18) is the transformation into quasi-classical form of the original definition of the canonical ensemble partition function

\[
Q_{NVT} \equiv \sum_{\Gamma} \exp(-\mathcal{H}(\Gamma)/kT),
\]

where \( \Gamma \) denotes a point in phase space, i.e. a state of the system. The probability density function that the system is in state \( \Gamma \) characterized by the Hamiltonian \( \mathcal{H}(\Gamma) \) is given by the Boltzmann probability distribution

\[
\mathcal{P}_{NVT} \equiv \frac{\exp(-\mathcal{H}(\Gamma)/kT)}{Q_{NVT}}
\]

(6.20)

Therefore, the ensemble average of a property \( A \) is defined as

\[
\langle A \rangle_{NVT} \equiv \sum_{\Gamma} \mathcal{P}_{NVT} A(\Gamma)
\]

(6.21)

for a discrete collection of states. For a continuous set of states \( \{\Gamma\} \) the average is given by the integral

\[
\langle A \rangle_{NVT} = \int d\Gamma \mathcal{P}_{NVT} A(\Gamma).
\]

(6.22)

Equations (6.19)-(6.22) are transformed into quasi-classical form (Girifalco, 1973) to give equations (6.18) and (6.16); see Girifalco (1973) for details.

The form of the Hamiltonian of eq. (6.17) implies that \( \mathcal{H} \) is a separable function of the coordinates and the momenta of the \( N \) particles. Therefore, separating the integrals involving the momenta in eq. (6.16) gives

\[
\langle A \rangle_{NVT} = \frac{\int d\mathbf{r}^N \exp(-\Phi(\mathbf{r}^N)/kT)A(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp(-\Phi(\mathbf{r}^N)/kT)},
\]

(6.23)

where \( A \) is a function of the coordinates of the particles only. Then the evaluation of \( \langle A \rangle_{NVT} \) is reduced to the calculation of two \( 3N \)-dimensional configurational integrals.

The Monte Carlo (MC) method (Kalos and Whitlock, 1986; Hammersley and Handscomb, 1964) has been routinely used for the calculation of configurational integrals in
in statistical mechanics (Binder, 1986). In the crude MC method the integrals are simply replaced by sums over the finite number $M$ of random configurations in the sample to give

$$
(A)_{NVT} = \frac{\sum_{m=1}^{M} A(m) \exp(-\Phi(m)/kT)}{\sum_{m=1}^{M} \exp(-\Phi(m)/kT)}.
$$

(6.24)

By sampling configurations at random from a chosen probability distribution $\mathcal{P}$ the integrals of eq. (6.23) can be estimated as

$$
(A)_{NVT} = \langle A\mathcal{P}_{NVT}/\mathcal{P} \rangle_{\text{trials}},
$$

(6.25)

where the average of the right side of eq. (6.25) is calculated over the generated configurations; each trial is an attempt to generate a new configuration. Equation (6.25) introduces the concept of importance sampling (Hammersley and Handscomb, 1964; Kalos and Whitlock, 1986). For most functions $A$ the integrand $A\mathcal{P}_{NVT}$ will be significant when $\mathcal{P}_{NVT}$ is significant. In these cases choosing $\mathcal{P} = \mathcal{P}_{NVT}$ should give a good estimate of the integral. Therefore, eq. (6.25) gives $\langle A \rangle_{NVT}$ as an unweighted average

$$
\langle A \rangle_{NVT} = \frac{1}{M} \sum_{m=1}^{M} A(M)
$$

(6.26)

over configurations in the sample.

An importance sampling method using $\mathcal{P} = \mathcal{P}_{NVT}$ was developed by Metropolis et al. (1953). The Metropolis algorithm generates a Markov chain of states of the system which has the limiting distribution of $\mathcal{P}_{NVT}$. A particle $i$ is selected in the simulation cell. The particle is displaced so that

$$
r_i' = r_i + d_{max} (2[\xi_1, \xi_2, \xi_3]^T - [1, 1, 1]^T),
$$

(6.27)

where $r_i'$ is the new particle position, $d_{max}$ is the maximum particle displacement, and $\xi_1, \xi_2, \xi_3$ are pseudo-random numbers uniformly distributed in the interval $[0,1]$. The new configuration is accepted or rejected according to the Metropolis criterion: if $\Delta E$ is the energy difference in the system due to the particle displacement the new configuration is
accepted if $\Delta E < 0$. If $\Delta E > 0$ the configuration is accepted for $\xi < \exp(-\Delta E/kT)$ and rejected for $\xi > \exp(-\Delta E/kT)$, where $\xi$ is a pseudo-random number in the interval $[0,1]$. Therefore, moves are accepted with a probability equal to

$$p_{\text{accept}} = \min[1, \exp(-\Delta E_{n-1,n}/kT)],$$

(6.28)

where $\Delta_{n-1,n}$ is the energy difference between the $n-1$ and $n$th configurations. In terms of averaging according to eq. (6.26) the system is considered to be in a new configuration for the purpose of averaging, whether or not the move is allowed.

6.2.2.2. Isothermal-Isobaric Ensemble Metropolis Monte Carlo

The isothermal-isobaric ($NpT$) ensemble is defined as a statistical ensemble of systems with fixed number of particles ($N$), pressure ($p$), and temperature ($T$). MC simulation in the ($NpT$) ensemble is a very useful tool for atomic modeling of solids because it allows the calculation of thermophysical properties which can be compared with experimental measurements taken at constant $p$ and $T$. The isothermal-isobaric partition function in quasi-classical form (Allen and Tildesley, 1989) is given by

$$Q_{NpT} = \frac{1}{N!} \frac{1}{h^3N} \frac{1}{V_0} \int dV \int d\mathbf{r}^N \int dp^N \exp\left[-\mathcal{H}(\mathbf{r}^N, p^N) + pV\right]/kT],$$

(6.29)

where $V_0$ is a volume used for normalization purposes.

The Metropolis algorithm is implemented by generating a Markov chain of states which has a limiting distribution proportional to

$$w = \exp\left[-(pV + \Phi(s))/kT + N \ln V\right],$$

(6.30)

where a set of scaled coordinates $s = (s_1, s_2, ..., s_N)$ is used with $s_i = L^{-1}r_i$ for a cubic simulation cell with cube edge $L = V^{1/3}$. A new state is generated by displacing an atom randomly and/or making a random volume change from $V_m$ to $V_n$ by

$$V_n = V_m + \delta V_{\text{max}}(2\xi - 1),$$

(6.31)
where $\delta V_{\max}$ is the maximum allowed change and $\xi$ is a pseudo-random number in the interval $[0,1]$. After the new state $n$ has been produced the quantity $\delta H$ is calculated as

$$
\delta H_{nm} = \delta \Phi_{nm} + p(V_n - V_m) - N k T \ln(V_n/V_m).
$$

(6.32)

The quantity $\delta H_{nm}$ is closely related to the enthalpy change in moving from state $m$ to state $n$. Moves are accepted with a probability equal to

$$
p_{\text{accept}} = \min[1, \exp(-\Delta H_{nm}/kT)],
$$

(6.33)

The pressure of the system is monitored during the $(NpT)$ MC simulation to examine if the calculated pressure is equal to the input pressure $p$ in the simulation. The pressure is calculated using the virial theorem (McQuarrie, 1976) as

$$
pV = N k T + \langle \mathcal{W} \rangle,
$$

(6.34)

where $\mathcal{W}$ is the internal virial (Allen and Tildesley, 1989) defined by

$$
\mathcal{W} \equiv \frac{1}{3} \sum_{i=1}^{N} r_i \cdot f_i,
$$

(6.35)

where $f_i$ is the interatomic force that acts on atom $i$. The interatomic force on atom $i$ is calculated from the gradient of the interatomic potential $\Phi$ with respect to the coordinates of particle $i$, i.e.

$$
f_i = -\nabla_{r_i} \Phi.
$$

(6.36)

The calculation of the interatomic forces for the SW interatomic potential and of the internal virial $\mathcal{W}$ are given in Section A.1 of the Appendix.

Cubic simulation cells are used and periodic boundary conditions are applied in the MC simulations (Hansen and McDonald, 1986; Allen and Tildesley, 1989). The minimum image convention (Allen and Tildesley, 1989) is used in the calculation of interatomic distances. Each simulation cell contains $N = 8n^3$, $n = 1, 2, 3, \ldots$ atoms because the conventional cubic unit cell in the diamond-cubic structure contains 8 atoms (Ashcroft and Mermin, 1976; Burns, 1985). The generation of configurations proceeds in repeated MC
sweeps, each sweep being a cycle starting from particle $i = 1$ and going to particle $i = N$. The maximum displacement $d_{max}$ and volume change $\delta V_{max}$ are adjusted after every 10 trials to maintain an acceptance ratio of 50% for trial moves. Volume adjustment is attempted after every MC sweep over the particles. The Verlet method (Verlet, 1967; Allen and Tildesley, 1989) is used to maintain a list of the neighbors of a particular atom, which is updated at intervals. The Verlet neighbor list is used to avoid the continuous examination of the interatomic separation between particles which are significantly further apart than the cutoff radius $r_c$ of the interatomic potential. The radius of the sphere centered on particle $i$ and which contains the particles in the neighbor list of $i$ is taken to have a radius $r_l/\sigma = a + 0.2$ in our simulations, where $a = r_c/\sigma$ is the cutoff scaled distance of the SW interatomic potential. The Verlet list is updated every 10 MC sweeps. For a system of $N = 216$ silicon atoms use of the so constructed and updated Verlet list speeds up the MC calculations by a factor of 3 with respect to the case where the list is not used.

6.2.2.3. Simulated Annealing

Simulated annealing is used for combinatorial optimization problems, i.e. to search for minima of functions of discrete variables. The method was originally developed by Kirkpatrick et al. (1983) who showed that a detailed analogy with annealing in solids provides a framework for combinatorial optimization of the properties of very large and complex systems. The simulated annealing method is based on the Metropolis algorithm. Although the simulated annealing method is based on statistical mechanical ideas it was originally applied to combinatorial optimization problems and later to continuous problems and especially to energy minimization of many-body systems (Van Laarhoven and Aarts, 1987).

Annealing in solids is the physical process in which a solid in a bath is heated up by increasing the temperature of the heat bath to a maximum value at which the particles of the solid arrange themselves in the liquid state followed by cooling through slowly lowering the temperature of the bath. In this way, all particles arrange themselves in the low energy ground state of a corresponding lattice provided that the maximum temperature is high
enough and that the cooling is carried out sufficiently slowly. At each temperature $T$ the solid is allowed to reach thermal equilibrium characterized by a probability of being in a state with energy $E$ given by the Boltzmann distribution. As $T$ decreases the Boltzmann distribution concentrates on the states with lowest energy and finally when $T \to 0$, only the minimum energy states have a nonzero probability of occurrence. If the solid is not allowed to reach thermal equilibrium at each $T$, defects can be frozen into the solid and metastable, amorphous structures can be reached.

Application of the simulated annealing method in statistical mechanical problems requires (i) the development of a cooling schedule, and (ii) implementation of the Metropolis algorithm at each annealing stage of temperature $T$. A cooling schedule consists of defining (i) the initial configuration and the maximum temperature value, (ii) the value of the minimum temperature where the algorithm is terminated, the system is frozen and no lowering in energy is accepted anymore, (iii) the length of the Markov chains in the configuration-generation process using Metropolis importance sampling, and (iv) a rule for lowering the temperature value. Both simple and elaborate cooling schedules have been reviewed by Van Laarhoven and Aarts (1987). In combinatorial problems a control parameter $(c)$ plays the role of the temperature and the cost function to be minimized plays the role of the system energy. Then, the simulated annealing algorithm is viewed as a sequence of Metropolis algorithms evaluated at a sequence of decreasing values of the control parameter.

The simulated annealing method avoids the main disadvantages of continuous nonlinear optimization methods (Gill et al., 1981) and combinatorial optimization methods (Van Laarhoven and Aarts, 1987). These disadvantages are the termination of the algorithm to a local minimum without any information on the amount by which this local minimum deviates from the global minimum and the dependence of the calculated local minimum on the initial configuration. Furthermore, in contrast to iterative improvement methods (Van Laarhoven and Aarts, 1987) simulated annealing can give a polynomial upper bound for the computational effort for some implementations of the algorithm.

The possibility of reaching the global energy minimum is guaranteed by the Metropolis algorithm. At a finite temperature there is always a nonzero probability of continuing with
a configuration of higher energy than the current configuration, i.e. a nonzero probability to climb uphill after reaching a local minimum. The minimum is approached by gradually decreasing this nonzero probability, which is achieved by lowering the temperature.

The main disadvantage of the simulated annealing method is its computational intensity. Using a few thousands of MC sweeps for every cooling stage and a total of 10 to 20 cooling stages for a system of a few hundreds or thousands of atoms constitutes a very lengthy MC simulation and makes the energy minimization routine extremely more expensive than any nonlinear optimization technique. Therefore, variations of the simulated annealing method have been used in the literature (Batra et al., 1987; Shoemaker et al., 1991; Ghaisas, 1991). In these variations the value of the maximum temperature is usually below the melting point.

The computationally most intensive energy minimization scheme used for our calculations is a variation of the simulated annealing method. The crystal is heated to $T_{\text{max}} = 1200$ K using an unrelaxed lattice as the initial configuration. The system is maintained at this temperature and zero pressure using the $(NpT)$ Metropolis MC method. After relaxation using very lengthy Markov chains, typically consisting of 10,000 to 20,000 MC sweeps, the volume is fixed to the average value at $p = 0$ and the system is mapped onto the local energy minimum configuration using the steepest descent method (Gill et al., 1981). The local energy minimum configuration is used as the initial configuration for the next cooling stage. The system is slowly cooled ($\Delta T \leq 100$ K) to 1 K using a total of 14 cooling stages.

The results of this elaborate minimization routine are reproduced to within an accuracy of $\pm 0.03$ eV following a much simpler minimization scheme. A very lengthy MC simulation, necessary for the calculations of thermodynamic properties, is performed at zero pressure and $T = 500$ K. Then the volume is set to the average value that corresponds to the zero pressure volume at $T = 0$ K. The appropriate rescaling of the particle coordinates is performed using the values of the scaled coordinates of the particles and the appropriate dimensions of the simulation cell. Minimization of the potential energy of the system following the steepest descent method is used as the final step in the calculation of the structure and the energy of the quenched configuration.
6.2.2.4. The Steepest Descent Method

The steepest descent method is an iterative method for \( n \)-dimensional nonlinear optimization which belongs to a class of methods that generate directions of search toward a local minimum without storing a matrix (Gill et al., 1981). Therefore, the value of the function \( \phi(x) \), where \( x \) is a \( n \)-dimensional vector, is minimized by iteratively updating the estimate of the minimum as

\[
x^{(k+1)} = x^{(k)} + a_k p^{(k)},
\]

where \( k \) denotes the number of the iteration. In the case of unconstrained optimization, the general algorithm consists of three steps: (i) computation of the nonzero vector \( p^{(k)} \) defined as the search direction, (ii) computation of the positive scalar \( a_k \) defined as the step length, and (iii) application of a convergence criterion to terminate the algorithm. The search direction is a descent direction if and only if

\[
\phi(x^{(k)} + a_k p^{(k)}) < \phi(x^{(k)}).
\]

Expansion of the left side of inequality (6.38) in a Taylor series and truncation of the series after the first-order term shows that the descent criterion is satisfied for search directions such that \( g^{(k)} \cdot p_k < 0 \), where \( g^{(k)} \) is the gradient of \( \phi \) evaluated at \( x^{(k)} \).

In the steepest descent algorithm the search direction is simply

\[
p^{(k)} = -g^{(k)} \equiv -\nabla \phi(x^{(k)}).
\]

The steepest descent is an efficient algorithm if we are away from the minimum and does not require the computation of the Hessian matrix of \( \phi(x) \). However, the method exhibits linear convergence, expressed in a vector norm \( \| \cdot \| \) as

\[
\|x^{(k+1)} - x_{\text{min}}\| \leq C\|x^{(k)} - x_{\text{min}}\|,
\]

where \( x_{\text{min}} \) is the exact solution that gives the local minimum value, and therefore converges very slowly in the long run. The linear convergence of the steepest descent algorithm is demonstrated in a trivial case of two-dimensional minimization where the exact solution
is known. In particular, the Rosenbrock function \( f(x, y) = 100(y - x^2)^2 + (1 - x)^2 \) (Gill et al., 1981), which exhibits a minimum value of zero at \( x = y = 1 \), is minimized using the steepest descent algorithm. The convergence plot is presented in Fig. 6.5 where the Euclidean norm is used as the norm in eq. (6.40). Using \( x_0 = y_0 = 1/2 \) as the initial guess the algorithm converges to the exact solution after 4135 iterations.

The step length \( a_k \) is calculated using line search algorithms. Our calculations use a minor modification of a method devised by Powell (Dahlquist and Björck, 1974). The algorithm is based on the use of a quadratic interpolation polynomial defined by three points \((\alpha, \psi(\alpha)), (\beta, \psi(\beta)), (\gamma, \psi(\gamma))\), where \( \psi(s) = \phi(x^{(k)} + s \mathbf{p}^{(k)}) \).

The correct calculation of the interatomic forces in our simulations using the SW atomic model is tested by energy minimization toward a known minimum. The descent direction defined by eq. (6.39) is the vector of the interatomic forces on all the particles in the system. The energy minimum for a silicon crystal described by the SW atomic model at a density \( \rho \sigma^3 = 0.4832 \) is \( \Phi/N \epsilon = -1.992346 \) and corresponds to a diamond-cubic crystal (Stillinger and Weber, 1985). The steepest descent algorithm should converge to this value of minimum energy for the appropriate initial configuration. The atomic configuration is a system of particles slightly perturbed from the ideal diamond lattice sites in a random way by using a maximum atomic displacement that is typically used in a MC simulation and a pseudo-random number generator. The convergence of the algorithm to the exact value after 24 iterations is demonstrated in Fig. 6.6.

The convergence criterion that is used for the termination of the steepest descent algorithm is

\[
|\Phi^{(k+1)}/\epsilon - \Phi^{(k)}/\epsilon| < 10^{-4}, \tag{6.41a}
\]

or

\[
||x^{(k+1)} - x^{(k)}|| < 10^{-4}, \tag{6.41b}
\]

where \( \Phi^{(k)} \) is the potential energy at the \( k \)th iteration, \( x \) is the \( 3N \)-dimensional vector of the three cartesian coordinates of all the \( N \) particles and \( ||\cdot|| \) is the Euclidean or \( L_2 \) norm.

Other methods of nonlinear optimization that have been used in lattice statics simulations are Newton's method (Wolf, 1984), the conjugate gradient method (Baskes et
Figure 6.5. Convergence of the steepest descent algorithm for the minimization of the Rosenbrock function.
Figure 6.6. Convergence of the steepest descent algorithm for the minimization of the potential energy of a silicon structure slightly perturbed from the ideal diamond lattice configuration. The convergence is demonstrated by (a) the Euclidean norm of the displacements of the atomic coordinates, and (b) the reduced potential energy of the crystal as a function of the number of steepest descent iterations.
al., 1989) and matrix-update methods (Gill et al., 1981) such as the method of Davidon, Fletcher and Powell (Sabochick et al., 1988).

6.2.3. Calculation of the Temperature Dependence of Thermodynamic Properties

The calculation of the temperature dependence of thermodynamic properties based on atomistic simulations is a straightforward task. Monte Carlo or molecular dynamics (MD) simulations are carried out at several temperatures and thermodynamic properties are computed as equilibrium ensemble averages or time averages (Hansen and McDonald, 1986; Allen and Tildesley, 1989) following statistical mechanical definitions. However, the computational effort required for accurate calculations over the temperature range $0 \leq T \leq T_m$ makes this calculation impractical in many studies of homogeneous and inhomogeneous solids described by many-body potentials.

The enormous computational intensity for the calculation of the temperature dependence of thermodynamic properties is circumvented by analyses of the energy or the enthalpy distribution functions obtained from a Monte Carlo simulation at a single temperature (Phillpot and Rickman, 1991). Cumulant analysis of the enthalpy distribution function (Rickman and Phillpot, 1991; Phillpot and Rickman, 1991) is used in our calculations of enthalpies, volumes, and Gibbs free energies of perfect and imperfect crystals as functions of temperature. The major advantage of this methodology is that a single $(NpT)$ MC simulation at a specific thermodynamic reference state $(p, T) = (p_{\text{ref}}, T_{\text{ref}})$ used as a reference state is required to determine the differences of thermodynamic properties at a temperature $T$ with respect to the values at $T = T_{\text{ref}}$.

The cumulant analysis of the enthalpy distribution function was recently developed by Phillpot and Rickman (1991) as a methodology for calculations of Gibbs free energies in homogeneous and inhomogeneous solids. The Gibbs free energy is defined (McQuarrie, 1976) as

$$G \equiv -kT \ln Q_{NpT},$$  \hspace{1cm} (6.42)

where the isothermal-isobaric partition function $Q_{NpT}$ is given by eq. (6.29). The defini-
tions of isothermal-isobaric ensemble averages are analogous to the definitions of canonical ensemble averages given by eqs. (6.16), (6.21) and (6.22). Using these definitions (Mc-Quarrie, 1976) gives

$$\Delta(\beta G) \equiv \beta G - \beta_{\text{ref}} G_{\text{ref}} = -\ln \langle \exp(\beta_{\text{ref}} - \beta) H \rangle_{\beta = \beta_{\text{ref}}},$$  \hspace{1cm} (6.43)

where $\beta \equiv 1/kT$, $\beta_{\text{ref}} \equiv 1/kT_{\text{ref}}$, $G_{\text{ref}}$ is the reference state Gibbs free energy, and $H$ is the enthalpy of the system and depends on the atomic configuration. In addition, any distribution is characterized by an infinite set of moments or equivalently by an infinite set of cumulants (Reichl, 1980) defined as

$$\langle \exp(\beta H) \rangle \equiv \exp \left[ \sum_{n=1}^{\infty} \frac{\beta^n}{n!} C_n(N, p, \beta) \right],$$  \hspace{1cm} (6.44)

where $C_n$ is the $n$th cumulant of the enthalpy distribution function. Combining eqs. (6.43) and (6.44) gives

$$\Delta(\beta G) \equiv \varepsilon G - \beta_{\text{ref}} G_{\text{ref}} = \sum_{n=1}^{\infty} \frac{(\beta - \beta_{\text{ref}})^n}{n!} (-1)^{n+1} C_n.$$  \hspace{1cm} (6.45)

The cumulants $C_n$ are defined by the ensemble averages

$$C_1 \equiv \langle H \rangle,$$  \hspace{1cm} (6.46a)

$$C_n \equiv \langle (H - \langle H \rangle)^n \rangle; \quad n > 1$$  \hspace{1cm} (6.46b)

which are calculated at the reference state. Equations (6.45) and (6.46) imply that a single MC simulation at the reference state is sufficient for the calculation of the free energy difference $\Delta(\beta G)$ for a range of temperatures.

The convergence properties of the infinite series in the cumulant expansion of eq. (6.45) were analyzed by Phillpot and Rickman (1991), who showed that truncation of the series keeping terms up to third order leads to accurate predictions of free energies up to the melting temperature $T_m$. However, the representation of $\Delta(\beta G)$ by a series expansion in powers of $(\beta - \beta_{\text{ref}})$ leads to divergence of the series as $T \to 0$ for a nonzero simulation.
temperature \( T_{\text{ref}} \). This is a result of the finiteness of \( G \) which leads to divergence of \( \beta G \) at temperatures close to zero (Rickman and Phillpot, 1991). The analysis has been extended by Rickman and Phillpot (1991) to provide expansions analogous to eq. (6.45) for other thermodynamic properties. Using the definition of the generalized cumulant expansion (Kubo, 1962) and introducing a joint moment generating function

\[
\phi(\lambda) \equiv \left\langle \exp \left[ - (\beta - \beta_{\text{ref}}) H + \lambda \delta A \right] \right\rangle, \tag{6.47}
\]

where

\[
\delta A \equiv A - \langle A \rangle \tag{6.48}
\]

gives (Rickman and Phillpot, 1991)

\[
\Delta A \equiv \langle A \rangle - \langle A \rangle_{\text{ref}} = \sum_{n=1}^{\infty} \frac{(\beta - \beta_{\text{ref}})^n}{n!} (-1)^n \left\langle \delta A \ (\delta H)^n \right\rangle, \tag{6.49}
\]

where

\[
\delta H \equiv H - \langle H \rangle. \tag{6.50}
\]

Rewriting eq. (6.49) in terms of the temperature \( T \) and then expanding \((1/T)^n\) in a Taylor series about \( T_{\text{ref}} \) expresses \( \Delta A \) in a Taylor series about \( T_{\text{ref}} \) as

\[
\Delta A = \sum_{n=1}^{\infty} \frac{1}{m!} \left[ \frac{\partial^m (\Delta A)}{\partial T^m} \right]_{T=T_{\text{ref}}} (T - T_{\text{ref}})^m, \tag{6.51}
\]

where

\[
\left[ \frac{\partial^m (\Delta A)}{\partial T^m} \right]_{T=T_{\text{ref}}} = \left( \frac{1}{T_{\text{ref}}} \right)^m m! \sum_{n=1}^{m} \frac{(-1)^{m-n}(m-n)}{n!} (n-1) \beta_{\text{ref}}^n K_n. \tag{6.52}
\]

The coefficients \( K_{n+1} \) are defined as the ensemble averages

\[
K_{n+1} \equiv \left\langle \delta A \ (\delta H)^n \right\rangle \tag{6.53}
\]

and are evaluated at the reference state.

The convergence properties of the Taylor series, eqs. (6.51) and (6.52), were analyzed by Rickman and Phillpot (1991) who showed that truncation of the series by keeping terms
up to second order leads to accurate calculations over the temperature range \(0 \leq T \leq T_m\).

This truncation requires accurate calculation of cumulants up to third order \((n + 1 = 3)\).

The basis of the truncation is the possibility to explore a wide range of the parameter space with the limited amount of information contained in the lower-order cumulants. Furthermore, the enthalpy distribution function for a large finite system is nearly Gaussian. Therefore, it is characterized by its first few cumulants only (Phillpot and Rickman, 1991). If the distribution indeed is Gaussian only the first two cumulants are nonzero (Phillpot and Rickman, 1991).

Our atomistic simulations at finite temperatures aim at the computationally efficient calculation of the temperature dependence of the enthalpy, volume and Gibbs free energy in perfect and imperfect crystalline silicon. Gibbs free energy differences from the reference state values are computed using eq. (6.45). Enthalpy and volume differences from the reference state values are computed using eqs. (6.51) - (6.53) for \(A = H\) and \(A = V\), respectively. For \(A = H\) eq. (6.53) gives

\[
K_{n+1} \equiv C_{n+1}, \quad n \geq 1. \tag{6.54}
\]

For \(A = V\) eq. (6.53) gives

\[
K_2 \equiv \left\langle (V - \langle V \rangle)(H - \langle H \rangle) \right\rangle, \tag{6.55a}
\]

\[
K_3 \equiv \left\langle (V - \langle V \rangle)(H - \langle H \rangle)^2 \right\rangle, \tag{6.55b}
\]

The ensemble averages of eqs. (6.54) and (6.55) are calculated by \((NpT)\) MC simulations at \(p_{\text{ref}} = 0\) and \(T_{\text{ref}} \leq 0.75\ T_m\). The accuracy of the method depends heavily on the accuracy of the calculation of the reference state Gibbs free energy. This method is described in Section 6.2.4. The reference values for the enthalpy and volume are computed directly from the simulation at \(T = T_{\text{ref}}\).

The enthalpy of a system for a specific atomic configuration is written as

\[
H = K + \Phi + pV, \tag{6.56}
\]

where \(K\) is the kinetic energy and \(\Phi\) is the potential energy of this configuration. Using the equipartition principle (McQuarrie, 1976; Allen and Tildesley, 1989) for the kinetic
energy contribution to the energy of a system of \( N \) atoms, expressions for the fluctuations in the kinetic energy in the canonical ensemble (Allen and Tildesley, 1989) and the transformations between ensembles gives

\[
\langle K \rangle = \frac{3}{2} N kT, \quad (6.57a)
\]

\[
\langle K^2 \rangle = (3N/2) (3N/2 + 1) (kT)^2, \quad (6.57b)
\]

and

\[
\langle K^3 \rangle = (3N/2) (3N/2 + 1) (3N/2 + 2) (kT)^2. \quad (6.57c)
\]

Using eqs. (6.57) and applying the definitions of the cumulants of the enthalpy distribution function, eqs. (6.46), at \( p = 0 \) gives

\[
C_1 = \frac{3}{2} N kT + \langle \Phi \rangle, \quad (6.58a)
\]

\[
C_2 = \frac{3}{2} N (kT)^2 + \langle (\Phi - \langle \Phi \rangle)^2 \rangle, \quad (6.58b)
\]

\[
C_3 = 3N (kT)^3 + \langle (\Phi - \langle \Phi \rangle)^3 \rangle. \quad (6.58c)
\]

Therefore, accurate computation of the cumulants is equivalent to the accurate computation of the cumulants of the interatomic potential, according to eqs. (6.58), by MC simulation at the reference state. In general, the second cumulant is directly related to the specific heat of the system under constant pressure \( C_p \) by

\[
C_p \equiv \left( \frac{\partial H}{\partial T} \right)_p = \frac{C_2}{kT^2}. \quad (6.59)
\]

Following a similar analysis and using that kinetic energy and volume are uncorrelated simplifies the expressions of the cumulants \( K_2 \) and \( K_3 \) in eqs. (6.55) as

\[
K_2 \equiv \left\langle (V - \langle V \rangle)(\Phi - \langle \Phi \rangle) \right\rangle, \quad (6.60a)
\]

\[
K_3 \equiv \left\langle (V - \langle V \rangle)(\Phi - \langle \Phi \rangle)^2 \right\rangle. \quad (6.60b)
\]
In general, the cumulant $K_2$ is directly related to the coefficient of thermal expansion $\alpha_p$ by

$$\alpha_p \equiv \frac{1}{\langle V \rangle} \left( \frac{\partial \langle V \rangle}{\partial T} \right)_p = \frac{K_2}{\langle V \rangle kT^2},$$  \hspace{1cm} (6.61)

where the average value $\langle V \rangle$ is used to denote that the volume is represented by the average value computed from the $(NpT)$ MC simulation.

The approximate expressions for the temperature dependence of the enthalpy and the volume using the expansion of eq. (6.51) up to $m = 2$ are

$$\langle H \rangle_T \simeq C_1 + \frac{C_2}{kT_{ref}^2} (T - T_{ref}) + \left[ \frac{C_3}{2k^2T_{ref}^4} - \frac{C_2}{kT_{ref}^3} \right] (T - T_{ref})^2,$$ \hspace{1cm} (6.62a)

and

$$\langle V \rangle_T \simeq \langle V \rangle_{ref} + \frac{K_2}{kT_{ref}^2} (T - T_{ref}) + \left[ \frac{K_3}{2k^2T_{ref}^4} - \frac{K_2}{kT_{ref}^3} \right] (T - T_{ref})^2,$$ \hspace{1cm} (6.62b)

Equations (6.62) are used to calculate the temperature dependence of the enthalpy and the volume of the systems of interest where the values of the cumulants $C_n$ and $K_n$ are calculated from our simulations, according to eqs. (6.58) and (6.60).

The success of the cumulant analysis of the enthalpy distribution function in predicting the temperature dependence of thermodynamic properties was originally demonstrated for Lennard-Jones solids (Phillpot and Rickman, 1990, 1991; Rickman and Phillpot, 1991). The approach was successful both for perfect single crystals and for polycrystals, i.e. inhomogeneous solids that contain grain boundaries. The accuracy of the cumulant analysis of the enthalpy distribution function was demonstrated by comparison with results from many MC simulations at different temperatures. The technique is estimated to be three to six times more computationally efficient than a temperature-integration scheme for free energy calculations (Phillpot and Rickman, 1991). Finally, the effects of the distribution of local strain on the free energy which should be taken into account together with the temperature changes were found to be very small even in the case of systems with large-scale defects (Phillpot and Rickman, 1990, 1991). The above results validate the application of
the cumulant analysis for the accurate and computationally efficient study of systems with point defects.

6.2.4. Free Energy Calculations

Lattice dynamics techniques within the harmonic or the quasi-harmonic approximations have been used traditionally for the calculation of free energies of homogeneous solids (Jacucci and Quirke, 1982; Jones and March, 1985). The disadvantages of the lattice dynamics methods at high temperatures or for inhomogeneous solids is that the harmonic approximation breaks down due to the increased anharmonic effects; under these conditions computationally intensive simulation techniques are required. Some of these techniques are the histogram method (Jacucci and Quirke, 1982; Phillpot and Rickman, 1991), the temperature integration method (Broughton and Li, 1987; Phillpot and Rickman, 1991), the method of energy sampling (Allen and Tildesley, 1989), the method of Frenkel and Ladd (1984) and the method of Broughton and Li (1987).

The method that we use for Gibbs free energy calculations is based on the cumulant analysis of the enthalpy distribution function, as described in Section 6.2.3. However, the cumulant analysis gives only free energy differences between two thermodynamic states of the system at the same pressure, one of which is used as the reference state. For calculations of absolute values of Gibbs free energies the cumulant analysis needs to be combined with a computationally efficient method for calculation of the Gibbs free energy at the reference state \( T = T_{\text{ref}} \). Unfortunately, the reference state Gibbs free energy of the system can not be calculated directly by the \((NpT)\) ensemble MC simulation used for the calculation of the reference state enthalpy and volume.

Recent atomistic simulations (Lutsko et al., 1988; LeSar et al., 1989) have proved that free energy calculations based on the theory of the harmonic crystal (Ashcroft and Mermin, 1976) give accurate predictions of the free energy of homogeneous solids and solids that contain point defects for temperatures up to 75% of the melting temperature \( T_m \). These calculations are based on the quasi-harmonic approximation (QHA) (Lutsko et al., 1988), and the local harmonic approximation (LHA) (LeSar et al., 1989). The results
of both methods have been compared with the results of exact methods for calculation of the free energy (Frenkel and Ladd, 1984). The use of two relatively short MD or MC simulations and of a single lattice statics simulation for the implementation of the QHA and the LHA, respectively, underline the tremendous computational advantages of these techniques. Alternatively, in the method of Frenkel and Ladd (1984) a series of new potentials is formed from weighted linear combinations of the interatomic potential of the solid and the interatomic potential of a reference solid, with the latter given by an Einstein potential with arbitrary force constant (Ashcroft and Mermin, 1976). The free energy of the solid is calculated with respect to that of the Einstein solid from a suitable integral of the energies associated with these new potentials (Frenkel and Ladd, 1984). Accurate calculation of the integral requires at least ten MD or MC simulations to provide a sufficient number of values for the integrand.

The basic assumption of the harmonic approximation (Ashcroft and Mermin, 1976) is that the atomic displacements in the crystal are very small from the equilibrium positions, which are the static lattice sites. Therefore, the potential energy of the crystal may be expanded in a Taylor series about the atomic coordinates of the equilibrium configuration. The linear term in the series vanishes because the gradient of the interatomic potential represents the force on each atom by all the other atoms, which must be zero when the atoms are in the equilibrium positions. The first nonvanishing correction to the equilibrium potential energy is given by the quadratic term in the series. In the harmonic approximation only the quadratic or harmonic term is retained and the potential energy is written as

\[ \Phi = \Phi_0 + \Phi_{\text{harm}}, \]  

(6.63)

where \( \Phi_0 \) is the potential energy of the system in the equilibrium configuration, and \( \Phi_{\text{harm}} \) is the quadratic correction.

The QHA is based on the assumption that the effect of the higher order terms in the series, which lead to thermal expansion, is taken into account by choosing the values of the lattice spacing in the configuration about which the expansion is made, so that this choice matches the equilibrium value at the desired spacing (Jacucci and Quirke, 1982). Therefore, the anharmonic corrections to the harmonic potential are incorporated partially
by readjusting the volume of the solid based on atomistic simulation using the appropriate interatomic potential. The quadratic correction in eq. (6.63) is written as the bilinear form

$$\Phi_{\text{harm}} = \frac{1}{2} u^T \cdot D \cdot u,$$  \hspace{1cm} (6.64)

where \( u \) is the \( 3N \)-dimensional vector of the displacements of the \( N \) particles of the system from their equilibrium positions and \( D \) is the dynamical matrix (Ashcroft and Mermin, 1976). The dynamical matrix is defined as the Hessian of the potential energy \( \Phi \) evaluated at the equilibrium configuration. Therefore, the elements of the \( 3N \times 3N \) dynamical matrix are

$$D_{ij,\alpha\beta} \equiv \left[ \frac{\partial^2 \Phi}{\partial r_{i,\alpha} \partial r_{j,\beta}} \right]_{r_{(o)}},$$  \hspace{1cm} (6.65)

where the indices \( i \) and \( j \) refer to the \( i \)th and \( j \)th particles, \( \alpha \) and \( \beta \) are cartesian indices, and the subscript outside the bracket denotes the equilibrium configuration.

In the classical limit the Helmholtz free energy of the harmonic crystal is computed from the canonical partition function (McQuarrie, 1976; Girifalco, 1973) as

$$F_{\text{harm}} \equiv -kT \ln Q_{NTV} = \Phi_0 + \frac{kT}{2} \ln \left[ \left( \frac{\Lambda^2}{2\pi kT} \right)^{3N} \prod_{i=1}^{3N} \omega_i^2 \right],$$  \hspace{1cm} (6.66)

where \( \omega_i \) are the normal mode frequencies and \( \Lambda \) is the thermal wavelength (McQuarrie, 1976) defined as

$$\Lambda \equiv \left( \frac{\hbar^2}{2\pi mkT} \right)^{1/2}.$$  \hspace{1cm} (6.67)

The normal mode frequencies are related to the eigenvalues \( \{ \lambda_i \} \) of the dynamical matrix \( D \) by

$$\lambda_i = \omega_i^2.$$  \hspace{1cm} (6.68)

Therefore, eq. (6.66) is rewritten as

$$F_{\text{harm}} = \Phi_0 + \frac{kT}{2} \ln \left[ \det \left( \frac{\Lambda^2}{2\pi kT} D \right) \right].$$  \hspace{1cm} (6.69)
where the determinant of the dynamical matrix, \( \det(\mathbf{D}) \), is the product of the squares of the \( 3N \) eigenfrequencies \( \{\omega_i\} \), i.e. \( \det(\mathbf{D}) \equiv \prod_{i=1}^{3N} \omega_i^2 \).

In the QHA simulations of Lutsko et al. (1988) a molecular dynamics run is performed initially in which the system is equilibrated to the desired thermodynamic state \((p = 0, T)\) using the method of Parrinello and Rahman (1981). The walls are then fixed and the molecular dynamics method is employed using the Nose thermostat (Nose, 1984) to generate the canonical ensemble and compute the equilibrium configuration that corresponds to the volume \( V \) at \( p = 0 \) and temperature \( T \). In the simulations the system is equilibrated to the thermodynamic state \((p, T) = (0, T_{\text{ref}})\) using the \((NpT)\) ensemble MC simulation. The canonical ensemble is generated and the equilibrium configuration is computed using the \((NVT)\) ensemble MC simulation. Using eq. (6.69) and the equality of the Helmholtz and the Gibbs free energy at \( p = 0 \) gives the reference state Gibbs free energy as

\[
G_{\text{ref}} = \Phi_0 + \frac{kT_{\text{ref}}}{2} \ln \left[ \det \left( \frac{\Lambda^2}{2\pi kT_{\text{ref}}} \mathbf{D} \right) \right],
\]

(6.70)

By equation (6.65), \( \mathbf{D} \) is a symmetric matrix. In addition, eq. (6.68) guarantees that all the eigenvalues of \( \mathbf{D} \) are not negative. There are three modes of zero frequency associated with the three translational degrees of freedom of the crystal. We eliminate the zero-frequency modes reducing the number of degrees of freedom of the system to \( 3N - 3 \) by keeping one of the particles fixed during the canonical ensemble MC simulation and eliminating the corresponding rows and columns of the dynamical matrix in the evaluation of the determinant (Hoover et al., 1972); the reduced \((3N-3) \times (3N-3)\) dynamical matrix is positive definite. The determinant of the reduced dynamical matrix is calculated using Choleski's algorithm for LU decomposition (Dahlquist and Björck, 1974).

Analytical calculation of the elements of the dynamical matrix for the SW interatomic potential is a very tedious algebraic task. The two-body and the three-body contributions to the diagonal and off-diagonal elements of the matrix are calculated separately. The result is given in final form as

\[
D_{mn,\alpha,\beta} = \frac{\partial^2 f_2(r_{mn})}{\partial r_{m,\alpha} \partial r_{n,\beta}} + \sum_{k \neq m \neq n} \frac{\partial^2 f_3(r_m, r_n, r_k)}{\partial r_{m,\alpha} \partial r_{m,\beta}},
\]

(6.71)

284
\[ D_{m,m',\alpha,\beta} = \sum_{i \neq m} \frac{\partial^2 f_2(r_{im})}{\partial r_{m,\alpha} \partial r_{m,\beta}} + \frac{1}{2} \sum_{j \neq m} \sum_{k \neq j \neq m} \frac{\partial^2 f_3(r_m, r_j, r_k)}{\partial r_{m,\alpha} \partial r_{m,\beta}}. \] (6.72)

The evaluation of the matrix elements of eqs. (6.71) and (6.72) is given in the form of a FORTRAN subroutine in Section A.2 of the Appendix. The correctness of our analytical evaluation of the dynamical matrix is tested by comparison with two numerical calculations. In the first calculation the gradient of the potential is calculated analytically and the dynamical matrix is computed by numerical calculation of the Jacobian matrix of the gradient, i.e. the Hessian matrix. In the second numerical calculation both the gradient and the Jacobian of the gradient are computed numerically. In all the numerical calculations first derivatives are computed by finite difference approximations (Dahlquist and Björck, 1974) which require multiple calculations of the system energy that corresponds to a small perturbation of one particle from its equilibrium position. Displacements as small as \(10^{-8}a\), where \(a\) is the lattice parameter are employed in the numerical evaluations of derivatives.

The local-harmonic approximation (LHA) is a further simplification of the harmonic approximation where all the terms that couple vibrations of different atoms are neglected (LeSar et al., 1989). The Helmholtz free energy of the crystal in the classical limit is written as

\[ F = \Phi_0 + \frac{kT}{2} \sum_{i=1}^{N} \ln \left[ \det \left( \frac{\Lambda^2}{2\pi kT} D_i \right) \right], \] (6.73)

where \(D_i\) is the local \((3 \times 3)\) dynamical matrix of particle \(i\) evaluated at the equilibrium configuration. The matrix elements are defined by

\[ D_{i,\alpha\beta} \equiv \left. \frac{\partial^2 \Phi}{\partial r_{i,\alpha} \partial r_{i,\beta}} \right|_{r_{i,\alpha}}, \] (6.74)

where the subscript outside the bracket denotes the equilibrium position of particle \(i\). Neglecting the coupling of the vibrations of different atoms simplifies the free energy calculation to the calculation of the determinants of \(N\) \((3 \times 3)\) matrices. This is the main computational advantage of the LHA compared to the QHA, where the memory requirements for large \((3N \times 3N)\) matrices are significant.
In the LHA simulations of LeSar et al. (1989) the equilibrium atomic structure and corresponding thermodynamic properties are determined by minimizing the free energy given by eq. (6.73) at any temperature with respect to the atomic coordinates. Therefore, a single lattice statics simulation is required, where the free energy is minimized instead of the energy. In our LHA calculations, the equilibrium configuration is generated following exactly the same procedure that is employed in the QHA calculations. Then, the Gibbs free energy of the system at the reference state \((p, T) = (0, T_{ref})\) is given by

\[
G_{ref} = \Phi_0 + \frac{kT_{ref}}{2} \sum_{i=1}^{N} \ln \left[ \text{det} \left( \frac{\Lambda^2}{2\pi kT_{ref}} D_i \right) \right].
\]  

(6.75)

LeSar et al. (1989) used the LHA to calculate the free energy of copper crystals using the Morse potential. The Helmholtz free energy of formation of the vacancy was calculated by the LHA although anharmonic effects are important in the crystal that contains the defect. The error of the calculation of the free energy with respect to the result obtained by application of the method of Frenkel and Ladd is only 1.2% at \(T = 960\) K, which is approximately 75% of the melting temperature for copper. This result validates the combination of atomistic simulation with harmonic models in the calculation of free energies for crystals with point defects over the temperature range \(0 \leq T \leq 0.75 T_m\). Furthermore, it supports our approach for Gibbs free energy calculations by combining the more accurate QHA for reference state free energy calculations with \(T_{ref} < 0.75 T_m\) with the cumulant analysis of the enthalpy distribution function. Earlier approaches for free energy calculations of systems with point defects have been reviewed by Jacucci and Quirke (1982).

### 6.2.5. Point Defect Concentrations at Thermal Equilibrium

Expressions for the thermal equilibrium concentration of intrinsic point defects in a crystal are derived following a thermodynamic formulation (Girifalco, 1973; Lannoo and Bourgoin, 1982) which is valid in the limit of dilute point defect concentrations, i.e.

\[
\frac{n}{N_s} \ll 1,
\]  

(6.76)
where \( n \) is the number of point defects in the crystal and \( N_s \) is the number of the crystalline lattice sites. In the dilute limit and \( p = 0 \) the Gibbs free energy of a crystal which contains a distribution of point defects is written as

\[
G = G_0 + n(H^f - TS^f) - TS^{conf},
\]

(6.77)

where \( G_0 \) is the Gibbs free energy of the perfect crystal, \( H^f \) is the formation enthalpy of an isolated point defect, \( S^f \) is the vibrational entropy of formation associated with the disorder induced by lattice vibration due to the formation of an isolated point defect, and \( S^{conf} \) is the configurational entropy induced by the distribution of the \( n \) defects over the \( N_s \) possible sites.

The formation properties of isolated point defects \( H^f \) and \( S^f \) are defined as the changes in the thermodynamic properties of the perfect crystal due to the introduction of an isolated point defect in the lattice. The exact definitions used for the calculation of the formation properties from atomistic simulation are given in Section 6.3.2. The configurational entropy due to the disorder induced by the introduction of a distribution of \( n \) point defects in the lattice is defined by

\[
S^{conf} \equiv k \ln \Omega,
\]

(6.78)

where \( \Omega \) is the complexion number (Lannoo and Bourgoin, 1982) which reduces to the number of the distinct ways that the \( n \) point defects can be distributed over the \( N_s \) lattice sites given by

\[
\Omega \equiv \binom{N_s}{n} = \frac{N_s!}{(N_s - n)! \cdot n!}.
\]

(6.79)

The number of intrinsic point defects at equilibrium \( n^{eq} \) is calculated by minimizing the Gibbs free energy of the crystal, eq. (6.77), with respect to the number of point defects. Using Stirling’s approximation (McQuarrie, 1976) and eq. (6.76) gives

\[
n^{eq} = N_s \exp\left(-\frac{G^f}{kT}\right),
\]

(6.80)

where \( G^f \) is the change in the Gibbs free energy of the crystal due to the formation of a single isolated point defect and is defined by

\[
G^f \equiv H^f - TS^f.
\]

(6.81)
Therefore, the equilibrium atomic fraction of point defects is defined by

\[ x^{eq} \equiv \frac{n^{eq}}{N_s} = \exp\left(-\frac{G^f}{kT}\right). \tag{6.82} \]

Equation (6.81) is usually rewritten by decomposing the right-hand side into a Boltzmann factor that contains the formation enthalpy and a preexponential factor that contains the vibrational entropy of formation as

\[ x^{eq} = \exp\left(\frac{S^f}{k}\right) \exp\left(-\frac{H^f}{kT}\right). \tag{6.83} \]

No assumption on the temperature dependence of the formation properties \((H^f, G^f, S^f)\) has been made for the derivation of eqs. (6.82) and (6.83). In general, formation properties are temperature dependent (Stott, 1978) and this temperature dependence is computed from our atomistic simulations reported in Sections 6.3.2 and 6.3.3.

Equation (6.82) is valid if and only if the number of the internal degrees of freedom of the intrinsic point defect \(n_{i.d.o.f.}\) is equal to unity. For \(n_{i.d.o.f.} \neq 1\) eq. (6.82) is rewritten as

\[ x^{eq} = n_{i.d.o.f.} \exp\left(-\frac{G^f}{kT}\right). \tag{6.84} \]

For vacancies \(n_{i.d.o.f.}\) is always equal to one. This is not true for some self-interstitial configurations. For example, the (100) split or dumbbell self-interstitial in fcc metals can be oriented in three equivalent ways, i.e. along the [100], [010], and [001] crystallographic axes (Johnson, 1973). In such a case \(n_{i.d.o.f.} = 3\) and \(\Omega = 3^n(N\choose n)\). Following exactly the same derivation gives equation (6.84). Quantum mechanical effects such as spin degeneracy (Fahey et al., 1989) are not included in our analysis for \(n_{i.d.o.f.}\).

In our atomistic simulations the temperature dependence of the thermal equilibrium atomic fractions of vacancies and self-interstitials are computed according to eq. (6.82). The calculation of the Gibbs free energy of formation follows the atomistic simulation approach described in Sections 6.2.3 and 6.2.4.
6.2.6. Jump Rate Theory of Point Defect Migration

The jump frequency $w$ of a point defect from one equilibrium site to another equilibrium site of the crystal was described in Section 4 following the jump rate or transition state theory in solids, according to the original formulation by Vineyard (1957). Introducing the harmonic approximation in the calculation of the configurational integrals of eq. (4.3) where the system is at the equilibrium and the saddle point configuration, respectively, gives the simple expression

$$w = \tilde{\nu} \exp\left(-\frac{H^m}{kT}\right),$$  \hspace{1cm} (6.85)

where $\tilde{\nu}$ is an effective attempt frequency (Girifalco, 1973) and $H^m$ is the migration enthalpy of the point defect. The migration enthalpy of the point defect is defined as the difference in the enthalpy of the crystal between the two configurations defined where the migrating atom is at the equilibrium and the saddle point positions, respectively (Girifalco, 1973).

The effective attempt frequency $\tilde{\nu}$ is defined in the harmonic approximation as

$$\tilde{\nu} \equiv \frac{\prod_{i=1}^{3N} \nu_i}{\prod_{i=1}^{3N-1} \nu_i^*},$$  \hspace{1cm} (6.86)

where $\nu_i = \omega_i/2\pi, \nu_i^* = \omega_i^*/2\pi$, and $\omega_i$ and $\omega_i^*$ are the normal mode (angular) frequencies when the system is at the equilibrium and the saddle point configuration, respectively. The missing degree-of-freedom in the product in the denominator of eq. (6.86) corresponds to the fixed coordinate of the jumping particle in the direction of the jump. This fixed coordinate is the coordinate that corresponds to the midpoint between the two equilibrium sites (Viñals and Sekerka, 1988).

Equations (6.85) and (6.86) can be simplified further by using the high temperature approximation ($h\nu \ll kT$) for the entropy of a harmonic system (Girifalco, 1973) as given by

$$S = -k \sum_{i=1}^{3N} \ln\left(\frac{h\nu_i}{kT}\right).$$  \hspace{1cm} (6.87)
With eq. (6.87), eq. (6.85) reduces to

\[ w = \frac{kT}{h} \exp\left( -\frac{G^m}{kT} \right), \quad (6.88) \]

where \( G^m \) is the Gibbs free energy of migration. The Gibbs free energy of migration is defined by

\[ G^m \equiv H^m - TS^m = G^* - G, \quad (6.89) \]

where \( S^m \) is the migration entropy defined as the entropy difference between the saddle point and the equilibrium point configuration and \( G^* \) denotes the Gibbs free energy of the crystal at the saddle point configuration. It should be mentioned again that no assumption has been made on the temperature dependence of the migration properties \( (H^m, G^m, S^m) \). In general, these quantities are treated as temperature dependent in our calculations.

The diffusion coefficient of intrinsic point defects is calculated following the random walk theory of isotropic diffusion (Girifalco, 1973). According to random walk theory the diffusion coefficient in every cubic-type lattice is given by

\[ D = \frac{1}{6} d_0^2 w, \quad (6.90) \]

where \( d_0 \) is the jump distance, i.e. the distance between two nearest-neighbor equilibrium sites. In diamond-cubic lattices the nearest-neighbor distance is \( d_0 = a\sqrt{3}/4 \), where \( a \) is the lattice parameter. Combining eq. (6.90) with eqs. (6.85) and (6.88) gives two expressions for \( D \):

\[ D = \frac{1}{6} d_0^2 \nu \exp\left( -\frac{H^m}{kT} \right), \quad (6.91a) \]

and

\[ D = \frac{1}{6} d_0^2 \frac{kT}{h} \exp\left( -\frac{G^m}{kT} \right). \quad (6.91b) \]

The temperature dependence of the diffusion coefficient is computed from eqs. (6.91) and the results of our atomistic simulations for \( H^m \) and \( G^m \), and for the temperature dependence of the lattice parameter of the perfect crystal \( a \). The effective frequency \( \nu \) is computed according to eq. (6.86) using the calculated values of the determinants.

290
of the dynamical matrix at the equilibrium and the saddle point position, respectively. The validity of eq. (6.91b) for expressing the temperature dependence of the diffusion coefficients is tested by comparison with the result of eq. (6.91a) at $T = T_{\text{ref}}$.

It should be noted that the calculation of the jump frequency following jump rate theory is based on equilibrium statistical mechanics. The activated state is one of the states in the ensemble representing the complete crystal equilibrium, and in computing the jump frequency, the frequency is counted by which the system moves from one set of states to another in the ensemble. In applying this theory to diffusion in a concentration gradient, it is assumed that the deviation from equilibrium caused by the nonzero gradients is not sufficient to affect seriously the equations derived in an equilibrium basis (Girifalco, 1973). Finally, computation of the diffusion coefficient of point defects at high temperatures is possible by using the molecular dynamics method to calculate the mean square displacement $\langle \Delta r^2 \rangle$ of the migrating particle and applying Einstein’s equation

$$D = \frac{1}{6} \frac{\langle \Delta r^2 \rangle}{\Delta t}. \quad (6.92)$$

However, the length of the molecular dynamics runs for statistically significant results in solid systems makes the MD calculation impractical in most of the cases of interest here.

6.3. Equilibrium and Transport Properties of Intrinsic Point Defects in Silicon

Results are presented in this section for the ground-state and the finite-temperature properties of vacancies and self-interstitials in silicon computed from atomistic simulations using the methods presented in Section 6.2 and the SW interatomic potential. Calculations are carried out for perfect crystals and imperfect crystals which contain one isolated vacancy and one isolated interstitial, respectively. Calculations in perfect crystalline silicon provide the ground for testing our computational method by comparison with results obtained from computationally intensive approaches. The results for perfect silicon also are used for the calculation of point defect formation properties. The results for the formation and migration properties of vacancies and self-interstitials are compared with experimen-
tal data and are used to analyze mechanisms of point defect diffusion and to compute contributions to self-diffusion due to vacancies and self-interstitials.

6.3.1. Thermodynamic Properties of Perfect Crystalline Silicon

The temperature dependence of the thermodynamic properties of perfect crystalline silicon is calculated from MC simulations using cubic simulation cells that contain \( N = 64, 216, \) and 512 atoms. Periodic boundary conditions are applied to eliminate surface effects. The results from statistically significant calculations of ensemble averages do not depend on the size of the perfect crystalline system. In addition, the values of the reduced cumulants (\( C_1/N, C_2/N^2, C_3/N^3, K_2/N^2, K_3/N^3 \)) are accurately reproduced for all the system sizes used in the simulations. Results are presented for systems with \( N = 216 \). The simulation cell containing 216 atoms in an unrelaxed diamond-cubic lattice configuration is presented in Fig. 6.7.

Lengthy Monte Carlo simulations are carried out at the reference state. The pressure is set to zero and the temperature is \( T = T_{\text{ref}} = 500 \) K, so that

\[
T_{\text{ref}} = 0.297 \; T_m, \tag{6.93}
\]

i.e. the reference temperature is sufficiently smaller than 75% of the melting temperature and inaccuracies in the calculation of the reference-state Gibbs free energies are avoided. 160,000 MC sweeps are used in the \((NpT)\) MC simulations and 60,000 of these are used for equilibration and 100,000 for averaging. These lengthy Markov chains are appropriate for accurate calculations of third order cumulants. A total of 50,000 MC sweeps are sufficient for the accurate calculation of second order cumulants. The need for sufficient equilibration of the volume of the system before the calculation of the cumulants \( K_2 \) and \( K_3 \) explains the length of the equilibration stage. The results for the cumulants \( C_n \) and \( K_n \) are:

\[
C_1^* \equiv C_1/(N\epsilon) = -1.9399 \pm 9.987 \times 10^{-7},
\]

\[
C_2^* \equiv C_2/(N\epsilon)^2 = 1.2110 \times 10^{-3} \pm 5.155 \times 10^{-7},
\]

\[
C_3^* \equiv C_3/(N\epsilon)^3 = 5.150 \times 10^{-5} \pm 5.15 \times 10^{-7},
\]
Figure 6.7. Cubic simulation cell containing 216 silicon atoms in a diamond-cubic lattice configuration.
\[ K_2^* \equiv K_2/(N^2\epsilon\sigma^3) = 2.5381 \times 10^{-4} \pm 1.888 \times 10^{-6}, \]
\[ K_3^* \equiv K_3/(N^3\epsilon^2\sigma^3) = 8.680 \times 10^{-6} \pm 1.213 \times 10^{-6}, \]
i.e. the relative errors in the calculations of \((C_1^*, C_2^*, K_2^*, C_3^*, K_3^*)\) are 5.15 \times 10^{-5}\%, 0.0043\%, 0.744\%, 1\% and 13.97\%, respectively. For the MC simulations in the canonical ensemble 60,000 MC sweeps are typically used. Markov chains of the same length were retained in the study of the imperfect crystalline systems with point defects.

Shorter \((NpT)\) and \((NVT)\) MC simulations also are carried out at temperatures of 250, 750, 1000, 1250, and 1500 K. At \(T = 0\) K only lattice sums calculations are involved. The simulations at temperatures different from the reference temperature are needed for calculations of ensemble averages for the enthalpy and the volume and for free energy calculations to be used for comparison with the results of the cumulant analysis. In all these simulations 60,000 MC sweeps are used including 10,000 MC sweeps for equilibration.

The success of the cumulant analysis of the enthalpy distribution function, eq. (6.62a), in predicting the temperature dependence of the enthalpy of the perfect crystal is demonstrated in Fig. 6.8. The analysis is based on the results of our reference-state calculations. Furthermore, the predictions of the cumulant analysis fall almost exactly on the curve computed from the polynomial correlation for \(E(T)\) derived from the MD simulations of Broughton and Li (1987), which is equal to \(H(T)\) at \(p = 0\).

The results of our calculations for the temperature dependence of the Gibbs free energy are presented in Fig. 6.9. The results of the cumulant analysis are compared with calculations at several temperatures based on the QHA and the LHA, and with the results of Broughton and Li (1987). Two sets of results are presented. The first is for the reduced Gibbs free energy per particle \(G/(N\epsilon)\) and the second for the quantity \(G/(N\epsilon) - g_0\), where
\[ g_0 \equiv \frac{3kT}{2\epsilon} \ln \left( \frac{\Lambda^2 \epsilon}{2\pi kT \sigma^2} \right), \quad (6.94) \]
is constant at constant temperature.

The results of Broughton and Li (1987) are derived by following a temperature integration scheme, characterized by its computational intensity. The Helmholtz free energy,
Figure 6.8. Comparison of the results of cumulant analysis with results from MC simulations for the temperature dependence of the enthalpy of a perfect silicon crystal.
Figure 6.9. Comparison of the results for the temperature dependence of the Gibbs free energy of a perfect silicon crystal of cumulant analysis with the results of the quasi-harmonic approximation (QHA), the local-harmonic approximation (LHA), and the temperature-integration scheme of Broughton and Li (1987).
equal to the Gibbs free energy at zero pressure, is given by

\[
F(T) = F_H(T) - T \int_0^T \frac{E(\tau) - E_H(\tau)}{\tau^2} \, d\tau. 
\]  

(6.95)

The free energy of the harmonic crystal \( F_H(T) \) was obtained by diagonalizing the dynamical matrix of a static lattice having the volume of the system at \( p = 0 \) and temperature \( T \). These results are considered exact and provide an excellent basis for comparison.

The success of the cumulant analysis in predicting Gibbs free energies at high temperatures is demonstrated in Table 6.2. For example, at \( T = 1000 \) K the relative error in \( G/(N\epsilon) - g_0 \) is 0.015% with respect to the results of Broughton and Li. The relative error in general is of the order of 0.1%. The validity of the QHA and the LHA at \( T < 0.75 \ T_m \) also is demonstrated. The relative error of the QHA and the LHA calculations of free energy at \( T = 0.74 \ T_m \) is 0.35% and 1.2%, respectively, with respect to the results of Broughton and Li (1987). Therefore, the approximate simulation techniques based on the harmonic approximation give accurate predictions for temperatures lower than 75% of the melting temperature. At \( T = 500 \) K, i.e. at the reference temperature for our calculations using the cumulant analysis, the relative error of the QHA result, which is our simulation value for the reference state Gibbs free energy, is 0.12% with respect to the corresponding result of Broughton and Li (1987). The divergence of the predictions of the cumulant analysis as \( T \to 0 \) is expected, as discussed in Section 6.2.3.

The results of the calculations are presented in Fig. 6.10 for \( \rho = \rho(T) = N/\langle V \rangle \) of the perfect crystal. The results from the \((NpT)\) MC simulations fall exactly on the curve that follows the polynomial expression for \( \rho(T) \) derived from the MD results of Broughton and Li (1987). The results of the cumulant analysis keeping only the linear term in the truncated expression, eq. (6.62b), are in good agreement with the MC results over the temperature range \( 0 \leq T \leq 750K \). Using the complete expression of eq. (6.62b) significantly improves the accuracy of the results, which are in excellent agreement with the MC results for \( 250 \) K \( \leq T \leq 1250 \) K.
Figure 6.10. Comparison of the results of cumulant analysis for $n = 2$ and $n = 3$ with results from MC simulations for the temperature dependence of the density of a perfect silicon crystal.
<table>
<thead>
<tr>
<th>T(K)</th>
<th>T/Tm</th>
<th>G/N\epsilon - g_0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Broughton and Li (1987)</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>-2.0000</td>
</tr>
<tr>
<td>250</td>
<td>0.1485</td>
<td>-1.8147</td>
</tr>
<tr>
<td>500</td>
<td>0.2971</td>
<td>-1.6712</td>
</tr>
<tr>
<td>750</td>
<td>0.4456</td>
<td>-1.5437</td>
</tr>
<tr>
<td>1000</td>
<td>0.5942</td>
<td>-1.4266</td>
</tr>
<tr>
<td>1250</td>
<td>0.7427</td>
<td>-1.3176</td>
</tr>
<tr>
<td>1500</td>
<td>0.8913</td>
<td>-1.2150</td>
</tr>
</tbody>
</table>

Table 6.2 Comparison of the predictions for the Gibbs free energy of the perfect crystal using the QHA, the LHA, and the cumulant analysis with the predictions using the temperature-integration scheme of Broughton and Li (1987).

6.3.2 Formation Properties of Extended Point Defects

Extended vacancies or self-interstitials are generated simply by removing one atom from the periodic simulation cell or by adding one atom to the simulation cell, respectively. In addition, neither the position of the extra atom is kept fixed at a particular lattice site nor the neighbors of the vacancy are prohibited to enter some spherical volume that surrounds the vacant site. Calculations are carried out for ground state properties using the lattice statics simulations described in Section 6.2.2 and for finite temperature properties using the MC simulations and the cumulant analysis of the enthalpy distribution function.

Application of periodic boundary conditions generates an array of periodically repeated point defects which are separated by a distance equal to the edge of the simulation cell. The elastic interaction between point defects is relatively short-range, proportional to the inverse sixth power of their separation distance (Eshelby, 1976; Lannoou and Bourgoin, 1982). However, if small simulation cells are used the interaction may be significant leading to a poor simulation model of the isolated defect. Alternatively, the use of larger and larger simulation cells increases the computational effort significantly, because the amount of work scales as \( N^2 \) with the system size. Therefore, smallest possible simulation cell
should be used where further increase of the size of the system does not affect the results for the formation properties.

Lattice statics simulations are carried out for systems where the host perfect crystal is represented by \( N = 64,216 \) and 512 atoms. For this test of convergence with system size the energy of the system is minimized using the steepest descent quench after relaxing the system with MC simulation at \( p = 0 \) and \( T = 500 \) K. The formation energies of the point defects are defined as the differences in the energies of the relaxed imperfect crystals with respect to the energy of the relaxed perfect crystal scaled to the same number of atoms as the imperfect crystal. Therefore, the formation energies of the vacancy and the self-interstitial are given by

\[
E_V^f \equiv E(N - 1) - (N - 1) \left[ \frac{E(N)}{N} \right],
\]

(6.96a)

and

\[
E_I^f \equiv E(N + 1) - (N + 1) \left[ \frac{E(N)}{N} \right].
\]

(6.96b)

These definitions have been used repeatedly in the literature (Benedek et al., 1992). An alternative definition suggested by Kwok et al. (1984) gives results which are consistently in agreement to within 0.05 to 0.1 eV of the results form the definitions of eqs. (6.96). The latter definition takes into account that a vacancy is formed by putting the missing atom to the surface and that the interstitial is formed by removing one atom from the surface into the bulk (Kwok et al., 1984).

The convergence of the calculation with system size is demonstrated in Fig. 6.11. We choose the system size \( N = 216 \) for which the calculation converges to less than 0.03 eV for the vacancy and to less than 0.01 eV for the self-interstitial. This system size also allows for reasonably fast computations of the cumulants by the use of very long Markov chains in the MC simulations.

The temperature dependence of the enthalpy is presented in Fig. 6.12 of the system with the extended vacancy and the extended self-interstitial, respectively. The reference state used in the cumulant analysis is the same as the one used for the perfect crystal; \( T = T_{\text{ref}} = 500 \) K and \( p = 0 \). The accuracy of the cumulant analysis is demonstrated by comparison with the results of 14 \( (NpT) \) MC simulations at several temperatures over
Figure 6.11. Dependence of the formation energy of vacancies and self-interstitials computed using the Stillinger-Weber potential on the number of atoms in the simulation cell.
Figure 6.12. Temperature dependence of the enthalpy for a silicon crystal with (a) an extended vacancy and (b) an extended self-interstitial per simulation cell.
the range $1 \text{ K} \leq T \leq 1200 \text{ K}$. The local energy minima for each state also are shown by mapping the atomic configuration at the end of each MC run onto the local energy minimum configuration using the steepest descent quench. This series of MC runs and steepest descent quenches constitutes one energy minimization calculation as described in Section 6.2.2. Lengthy Markov chains of 20,000 MC sweeps are used at each temperature to allow for accurate averaging. The results of Fig. 6.12 prove that the cumulant analysis of the enthalpy distribution function gives accurate results for both the ground state and the high temperature properties of the imperfect systems. The formation energies computed for the vacancy and the self-interstitial are 2.65 eV and 3.63 eV, respectively. The cumulant analysis method is more computationally efficient by a factor of two compared to the sequence of MC runs and steepest descent quenches.

The distribution of the local energy minima that correspond to the 14 thermodynamic states of each imperfect crystal presented in Fig. 6.12 are plotted in Fig. 6.13 using an expanded scale to demonstrate the details in the results of the energy minimization procedure. The result for the formation energy of the extended vacancy obtained by the alternative minimization routine described in Section 6.2.2 is exactly the same, i.e. 2.65 eV. The corresponding result for the formation energy of the extended self-interstitial is a higher value only by 0.03 eV, i.e. $E_f^c = 3.66$ eV. These very small energy differences validate the use of the more computationally efficient minimization routine.

The atomic relaxations of the lattice caused by an extended vacancy and an extended self-interstitial are presented in Fig. 6.14. The extend interstitial is generated by putting an extra atom to a tetrahedral site close to the center of the simulation cell where all the other atoms occupy the lattice sites of an unrelaxed diamond-cubic lattice. The differences between the values $r_q$ of the quenched atomic positions and the distances $r_u$ of the unrelaxed atomic positions form the original vacant site and the original tetrahedral site, respectively, show the pronounced relaxations in the first and second nearest-neighbor shells. The significant nonuniformity of the relaxation also is evident within a given shell. In the case of the vacancy the relaxation is inward as expected, since the vacancy is a center of dilatation (Eshelby, 1976). The maximum displacement experienced by one of the nearest-neighbors of the original vacant site is slightly less than a jump distance and
Figure 6.13. Local energy minima that correspond to the atomic configuration at the end of a cooling stage as a function of the temperature of the cooling stage using a variation of the simulated annealing method for energy minimization of a silicon crystal that contains (a) a vacancy and (b) a self-interstitial per simulation cell.
Figure 6.14. Atomic relaxation in units of lattice parameter (a) within neighboring shells of (a) an extended vacancy, and (b) an extended self-interstitial in a silicon crystal.
emphasizes the strong anharmonic effects in the presence of the vacancy. In the case of the extended interstitial the relaxation is both inward and outward. The maximum displacement from the initial position is experienced by the extra atom and is approximately equal to 60% of the jump distance. This significant relaxation indicates again the strong anharmonic character of the system with the extended interstitial. However, in both cases simulations based on harmonic models capture the strong anharmonic effects at temperatures away from the melting point, because the interatomic potential accounts for the appropriate lattice spacing and produces the correct thermal expansion (LeSar et al., 1989; Phillpot and Rickman, 1991). Finally, the relaxation decays exponentially away from the defect center.

The enthalpy and Gibbs free energy of formation for the vacancy and the interstitial are defined as

\[
H_V^f \equiv H(N - 1) - (N - 1) \left[ \frac{H(N)}{N} \right], \tag{6.97a}
\]

\[
H_I^f \equiv H(N + 1) - (N + 1) \left[ \frac{H(N)}{N} \right], \tag{6.97b}
\]

\[
G_V^f \equiv G(N - 1) - (N - 1) \left[ \frac{G(N)}{N} \right], \tag{6.97c}
\]

\[
G_I^f \equiv G(N + 1) - (N + 1) \left[ \frac{G(N)}{N} \right], \tag{6.97d}
\]

in exactly the same manner as the formation energies. Then, the vibrational entropies of formation are

\[
S_V^f \equiv \frac{1}{T} \left( G_V^f - H_V^f \right), \tag{6.98a}
\]

and

\[
S_I^f \equiv \frac{1}{T} \left( G_I^f - H_I^f \right). \tag{6.98b}
\]

The temperature dependence of the enthalpies, Gibbs free energies and vibrational entropies of formation of point defects are presented in Fig. 6.15 as predicted by the cumulant analysis of the enthalpy distribution function. At the reference state, the formation enthalpies, free energies, and vibrational entropies for the vacancy and the self-interstitial are (2.61 eV, 2.52 eV, 2.2 \textit{k}), and (3.67 eV, 3.43 eV, 5.5 \textit{k}), respectively. The variation of the formation enthalpies with temperature is small for both defects. The small variation
Figure 6.15. Temperature dependence of the enthalpy ($H^f_X$), Gibbs free energy ($G^f_X$), and vibrational entropy ($S^f_X$) of formation of extended vacancies ($X = V$) and self-interstitials ($X = I$) in silicon.
of the vacancy formation enthalpy with temperature predicted by our calculations is in qualitative agreement with measurements from positron annihilation experiments (Dannefaer et al., 1986). The maximum errors in the calculations of formation enthalpies and Gibbs free energies at high temperatures are estimated as ±0.05 eV and ±0.2 eV respectively.

The temperature dependence of the density of the crystals with the vacancy and the self-interstitial predicted by cumulant analysis is presented in Fig. 6.16 together with the results from the 14 MC simulations for the range 1 K ≤ T ≤ 1200K. The results of the cumulant analysis keeping only the linear term in the truncated expansion for V(T), eq. (6.62b), are in good agreement with the MC results for 0 ≤ T ≤ 800 K and 100 K ≤ T ≤ 1000 K for the system with the vacancy and the self-interstitial, respectively. Using the complete expression of eq. (6.62b) improves the accuracy of the results for the system with the vacancy and gives very good agreement with the MC results for 200 K ≤ T ≤ 1200 K.

The relaxation and formation volumes of the vacancy and the self-interstitial are defined as

\[ \Delta V^\text{rel}_V \equiv V(N - 1) - V(N), \]  
\[ \Delta V^\text{rel}_I \equiv V(N + 1) - V(N), \]  
\[ V^f_V \equiv V(N - 1) - (N - 1) \left[ \frac{V(N)}{N} \right] = \Delta V^\text{rel}_V - \Omega_p, \]  
\[ V^f_I \equiv V(N + 1) - (N + 1) \left[ \frac{V(N)}{N} \right] = \Delta V^\text{rel}_I - \Omega_p, \]

where \( \Omega_p \) is the atomic volume of the perfect crystal \( V(N)/N \) and the volumes denote equilibrium ensemble averages computed from \( (NpT) \) MC simulations. The results are presented in Fig. 6.17 for the temperature dependence of the formation and relaxation volumes for the vacancy and the self-interstitial, respectively, predicted both by cumulant analysis and MC simulation. At the reference state the formation volumes of the vacancy and the interstitial are \(-0.74 \Omega_p \) and \(0.33 \Omega_p\), respectively, where \( \Omega_p \) is computed at the same thermodynamic state \( T = T_{ref} = 500 \) K, \( p = 0 \). The predicted variation with temperature is small, as in the case of the formation enthalpies.
Figure 6.16. Comparison of the results of cumulant analysis for $n = 2$ and $n = 3$ with results from MC simulations for the temperature dependence of the density, $\rho = \rho(T)$, of a silicon crystal that contains extended vacancies or extended self-interstitials.
Figure 6.17. Temperature dependence of the relaxation volume ($\Delta V_X^{\text{rel}}$) and the formation volume ($V_X^f$) of the extended vacancy ($X = V$) and the extended self-interstitial ($X = I$).
6.3.3 Formation Properties of Strictly Localized Point Defects

Calculations are performed for strictly localized point defect configurations of high symmetry for two reasons. The first reason is to compare the formation properties with the results obtained for the extended point defect configurations and to determine the configurations of lowest energy, which should be used to calculate the corresponding thermal equilibrium concentrations. The second reason is to compare our results for localized point defects to other atomistic simulations using empirical and semi-empirical interatomic potentials (Batra et al., 1987; Baskes et al., 1987; Ghaisas, 1991) and ab initio total energy calculations (Car et al., 1984, 1985) in order to reach more general conclusions on the ability of the SW potential to model point defects in silicon. These comparisons will help to validate quantitative comparisons with experimental data and clarify or suggest mechanisms and paths for point defect migration.

Six intrinsic point defect configurations are analyzed: the vacancy (V), the split-vacancy (SV), and the terahedral (I_T), the hexagonal (I_H), the bond-centered (I_B), and the [100] split or dumbbell (I_D self-interstitials. In all six cases the point defect is put close to the center of the cubic simulation cell and the initial configuration is the unrelaxed lattice. In the case of the vacancy one of the nearest neighbors of the original vacant site is held fixed during the simulation. In the case of the split-vacancy one of the nearest neighbors of the vacant site is put at the midpoint between its original position and the vacant site and is held fixed during the simulation. In the cases of the tetrahedral, the hexagonal, and the bond-centered interstitials the extra atom is put at the tetrahedral site, the hexagonal site, and the midpoint between two neighboring atoms, respectively, and is held fixed during the simulation. Finally, in the case of the dumbbell configuration the extra atom and the host lattice atom that form the dumbbell are both held fixed during the simulation.

In the \((NpT)\) MC simulations the scaled positions of the atoms that define the point defect configurations are held fixed with respect to the center of mass of the simulation cell, while the actual coordinates are adjusted following the volume adjustments that maintain \(p = 0\). In the \((NVT)\) MC simulations the actual positions and the center of mass
(Hoover et al., 1972) are held fixed. The same coordinates also are held fixed during the steepest descent quench which takes place under constant volume which is the average volume computed from the simulation at $p = 0$. The modification of the steepest descent algorithm in the case that some coordinates are held fixed is straightforward. Holding some particle coordinates fixed is the simplest linear equality constraint that can be posed to the minimization problem (Gill et al., 1981) and the steepest descent algorithm for the unconstrained problem remains essentially unchanged. The fixed coordinates do not participate in the update of the estimate for the minimum given by eq. (6.37).

The results of the calculations for the formation energies of the relaxed configurations and the average first-neighbor relaxation are presented in Table 6.3 together with the predictions of modified embedded-atom method (MEAM) of Baskes et al. (1989). The results for the formation energies of the strictly localized self-interstitial configurations predicted by the SW potential are in excellent agreement with the results of Batra et al. (1987) using the SW interatomic potential. Batra et al. (1987) reported the values of 4.95, 6.54, 5.61, and 5.26 eV for the formation energies of $I_T$, $I_H$, $I_B$, and $I_D$, respectively. The maximum difference is 0.12 eV for the dumbbell configuration resulting in a relative error of 2.2% with respect to the calculation of Batra et al. (1987). The agreement is considered excellent because different system sizes and different simulation techniques were used for the two calculations. Batra et al. (1987) used a rectangular simulation cell with $N = 800$ atoms for the perfect crystal, and the relaxation was carried out using the MD method at $(T, p) = (0, 0)$. In our calculation relaxation is carried out using the steepest descent quench after a lengthy $(NpT)$ MC simulation at zero pressure and $T = 500$ K. Our result of 2.66 eV for the formation energy of the vacancy disagrees with the value of 4.34 eV reported by Batra et al. (1987). This value is exactly the same as the formation energy of the unrelaxed configuration. The very small energy loss for the vacancy upon relaxation was recognized by Batra et al. (1987), who also carried out the relaxation using a variation of the simulated annealing method. The new minimization scheme suggested the existence of a manifold of energy minima resulting in different atomic positions around the vacancy. However, no further quantitative predictions were reported by Batra et al. (1987).
Table 6.3 Formation energies and first-neighbor relaxations of intrinsic point defects in silicon predicted by our calculations using the SW interatomic potential and by Baskes et al. (1989) using the modified embedded-atom method (MEAM).

Comparison of our results with the results of Baskes et al. (1989) for the energetics of point defects is interesting because of the very different conceptual foundations between the embedded-atom method (Daw and Baskes, 1984) and the SW atomic model. Interestingly both atomic models predict the same ordering of the energies for vacancy and interstitial formation as a function of the configuration of the defect. In addition, the corresponding formation energies are quantitatively similar to within 0.5 eV. Furthermore, the lower energies for vacancy formation than interstitial formation suggest that the vacancy mechanism is dominant for self-diffusion; this qualitative result follows from calculations with both atomic models. Moreover, comparison of the formation energies for the vacancy and the split-vacancy, which is expected to be the saddle point configuration for vacancy migration give migration energies of 0.43 and 0.37 eV predicted by the SW atomic model and the MEAM, respectively. These values are surprisingly close for two completely different atomic models. Finally, the predicted formation energies for the strictly localized self-interstitial configurations from both atomic models are in reasonable quantitative agreement with the results of the total energy calculations of Car et al. (1984, 1985). The ordering of the energies for interstitial formation from both atomic models are in agreement with the total energy calculations for the doubly charged state (I$^{2+}$) but in
disagreement with the ordering for the neutral charge state \((E_B^f < E_H^f < E_T^f)\) predicted by Car et al. (1984, 1985).

Comparison of our results with the MEAM predictions for the average first-neighbor relaxation also shows very good agreement for the cases of the split-vacancy and the tetrahedral, hexagonal and bond-centered self-interstitials. The prediction of both models for the relaxation also are in qualitative agreement with results from total energy calculations (Car et al., 1984), which give relaxations that range from being negligible for the T site, 0.1 Å for the H site and to 1.2 Å for the B site.

The atomic relaxations of the lattice caused by the strictly localized point defects are presented in detail in Figs. 6.18, 6.19, and 6.20 for the cases of the vacancy and the split-vacancy, the tetrahedral and the hexagonal self-interstitials, and the bond-centered and the dumbbell self-interstitials, respectively. The relaxation with respect to the position of the point defect is inward in the case of the vacancy and the split-vacancy and outward for all the self-interstitial configurations. Again the relaxation is expressed by the differences between the distances \(r_q\) of the quenched atomic positions and the distances \(r_u\) of the unrelaxed atomic positions from the point defect as a function of the unrelaxed distances \(r_u\). In all cases the relaxation is oscillatory and nonuniform within a given neighboring shell. Strong anharmonic effects arise from the relaxation around the vacancy and the bond-centered self-interstitial. The smallest lattice relaxation is observed around the high symmetry (four-fold coordination) tetrahedral interstitial. In all cases the relaxation decays exponentially away from the defect center. Our results for the atomic relaxations are in excellent agreement with the results of Batra et al. (1987) who also reported for the long-range relaxation a correlation of the form

\[
r_q - r_u = \xi \exp(-\beta r_u),
\]  

(6.100)

where \(\xi = 0.17, \beta = 2.5\) and the distances are given in units of the lattice parameter \(a\). In general, up to a distance \(r_u = a\), relaxation deviates from the above expression and exhibits oscillations with significant amplitudes. For \(r_u/a > 1\) the amplitudes of the oscillations become smaller and the relaxations decay exponentially according to (6.100).

The temperature dependence of the enthalpy, Gibbs free energy and vibrational
Figure 6.18. Atomic relaxation in units of lattice parameter (a) within neighboring shells of (a) a vacancy, and (b) a split-vacancy in a silicon crystal.
Figure 6.19. Atomic relaxation in units of lattice parameter \((a)\) within neighboring shells of (a) a tetrahedral, and (b) a hexagonal self-interstitial in a silicon crystal.
Figure 6.20. Atomic relaxation in units of lattice parameter \( a \) within neighboring shells of (a) a bond-centered, and (b) a dumbbell self-interstitial in a silicon crystal.
entropy of formation for all the strictly localized point defects are calculated using the cumulant analysis of the enthalpy distribution function. It is recognized that the fixed particle coordinates correspond to zero-frequency modes for the Gibbs free energy calculations at the reference state, which result in lowering the number of degrees-of-freedom of the crystal. This effect is taken into account by eliminating the corresponding rows and columns of the dynamical matrix in the evaluation of the determinant. The results are presented in Figs. 6.21, 6.22, and 6.23 for the vacancy and the split-vacancy, the tetrahedral and the hexagonal self-interstitials, and the bond-centered and dumbbell self-interstitials, respectively. The variation with temperature of the formation enthalpies and Gibbs free energies is small in all cases. The maximum errors in the calculation of formation enthalpies and Gibbs free energies at high temperatures are estimated as 0.04 eV and 0.2 eV, respectively.

The temperature dependences of the formation volumes of the vacancy and split-vacancy configurations, and for all the localized interstitial configurations are presented in Figs. 6.24 and 6.25, respectively. The computed variation of the formation volumes with temperature is much higher than the corresponding variation for extended point defect configurations. The predictions from our simulations for the enthalpy, Gibbs free energy, vibrational entropy and volume of formation of all the types of strictly localized point defect configurations at the reference state \( p = 0, T = 500 \text{ K} \) are given in Table 6.4.

<table>
<thead>
<tr>
<th>Point Defect</th>
<th>( H^f ) (eV)</th>
<th>( G^f ) (eV)</th>
<th>( S^f/k )</th>
<th>( V^f/\Omega_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>2.64</td>
<td>2.53</td>
<td>2.46</td>
<td>-0.73</td>
</tr>
<tr>
<td>SV</td>
<td>3.09</td>
<td>2.98</td>
<td>2.56</td>
<td>-0.70</td>
</tr>
<tr>
<td>IT</td>
<td>4.81</td>
<td>4.56</td>
<td>5.89</td>
<td>1.06</td>
</tr>
<tr>
<td>IH</td>
<td>6.57</td>
<td>6.28</td>
<td>6.82</td>
<td>0.65</td>
</tr>
<tr>
<td>IP</td>
<td>5.66</td>
<td>5.58</td>
<td>1.80</td>
<td>0.42</td>
</tr>
<tr>
<td>ID</td>
<td>5.34</td>
<td>5.27</td>
<td>1.74</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 6.4 Formation properties of strictly localized intrinsic point defect configurations computed at the reference state, \( p = 0 \) and \( T = 500 \text{ K} \).
Figure 6.21. Temperature dependence of the enthalpy ($H_X^f$), Gibbs free energy ($G_X^f$), and vibrational entropy ($S_X^f$) of formation of the vacancy ($X = V$) and the split-vacancy ($X = SV$) in silicon.
Figure 6.22. Temperature dependence of the enthalpy ($H_X^f$), Gibbs free energy ($G_X^f$), and vibrational entropy ($S_X^f$) of formation of the tetrahedral ($X = T$) and the hexagonal ($X = H$) self-interstitial in silicon.
Figure 6.23. Temperature dependence of the enthalpy ($H_X^f$) and the Gibbs free energy ($G_X^f$) of formation of the bond-centered ($X = B$) and the dumbell ($X = D$) self-interstitial in silicon.
Figure 6.24. Temperature dependence of the formation volume of the vacancy and the split-vacancy.
Figure 6.25. Temperature dependence of the formation volume of strictly localized self-interstitial configurations.
6.3.4. Thermal Equilibrium Concentrations

The results of the calculations for the temperature dependence of the formation Gibbs free energies of point defects are combined with eq. (6.82) to calculate the equilibrium atomic fractions $x^{eq}$ for vacancies and self-interstitials. The equilibrium concentrations per unit volume $c^{eq}$ are defined as

$$c^{eq} \equiv x^{eq} \frac{N_A \rho}{(AW)},$$

(6.101)

where $N_A = 6.023 \times 10^{23}$ atoms/mol is Avogadro’s number, $\rho$ is the density of the perfect crystal and $(AW)=28.086$ g/mol is the atomic weight of silicon. The results of the simulations for the temperature dependence of the density of the perfect crystal are used to carry out the calculation in eq. (6.101).

The ordering of the energy for vacancy and self-interstitial formation is the appropriate criterion that determines the stable configuration to be used in the calculation of the equilibrium concentration of vacancies $c_v^{eq}$ and self-interstitials $c_i^{eq}$, respectively. In both cases the extended point defects are the favorite configurations. In the case of the vacancy the formation energies of the extended configuration and the configuration with one of the nearest neighbors fixed are almost identical; the corresponding values of the formation energies are 2.65 and 2.66 eV, respectively. Alternatively, the extended interstitial has a formation energy that is 1.18 eV lower than the energy of the strictly localized interstitial configuration of lowest energy; i.e. the highly symmetric tetrahedral interstitial. Therefore, the calculated temperature dependence of the Gibbs free energies of formation of the extended point defect configurations are used in the calculation of the equilibrium atomic fractions.

The low energy of the extended self-interstitial configuration from our calculations support the extended interstitial model that was proposed to explain the high value of the preexponential factor in the coefficient of self-diffusion (Seeger and Chik, 1968). The results of the total energy calculations of Car et al. (1984, 1985) also support the extended interstitial model. The reason is the multiplicity of self-interstitial configurations of similar energy and the significant lattice relaxations that accompany some of these configurations.
(Car et al., 1984, 1985). Therefore, significant contributions to the vibrational and configurational entropy are expected to lead to a large preexponential factor for the self-diffusion coefficient. Our calculated average vibrational entropies of formation for the extended and tetrahedral interstitials are 5.5 \( k \) and 4.0 \( k \), respectively, for 500 K \( \leq T \leq 1500 \) K; these values are higher by 1.5 \( k \) for the extended defect.

The predictions of our calculations for the equilibrium concentrations of vacancies and self-interstitials are presented in Fig. 6.26 for 909 K \( < T < 1635 \) K. Systematic analysis of experimental data (Tan and Gösele, 1985) has shown that this is the temperature range where both vacancies and self-interstitials diffuse fast, resulting in similar contributions to the self-diffusion coefficient in silicon. All the experimental results are expressed in simple Arrhenius forms resulting in straight lines for the corresponding Arrhenius plots. The slight curvature observed in our simulation predictions is a result of the temperature dependent Gibbs free energies of formation predicted by our calculations.

The simulation results for the vacancy are compared with results from two different analyses of the positron annihilation data of Dannefaer et al. (1986) and the analysis of dopant diffusion data of Masters and Gorey (1978). The analysis of the positron data of Dannefaer et al. (1986) gives the expression

\[
c_{V}^{c} = 5 \times 10^{22} \exp(8.0 \pm 2.0) \exp\left[-\frac{(3.6 \pm 0.2) \text{ eV}}{kT}\right] \text{ cm}^{-3}, \tag{6.102}
\]

over the temperature range 1250 K \( < T < 1500 \) K. However, analysis of the same set of data by Van Vechten (1987) gives the expression

\[
c_{V}^{c} = 5 \times 10^{22} \exp(2.9) \exp\left(-\frac{2.6 \text{ eV}}{kT}\right) \text{ cm}^{-3}, \tag{6.103}
\]

The source of the differences in the analysis of the same set of data was discussed in Section 6.1. Interestingly enough the results of our simulations are intermediate between these two expressions. The simulation results give a curve that almost parallels the result of eq. (6.103) due to the similarity in the activation energy in eq. (6.103) and our results for the formation energies of vacancies. However, at temperatures close to 1500 K the simulation results are in reasonable agreement with the average values given in eq. (6.102). The
Figure 6.26. Temperature dependence of the thermal equilibrium concentration of (a) vacancies, and (b) self-interstitials. The simulation predictions are compared with results from analyses of positron annihilation and dopant diffusion data.
analysis by Masters and Gorey of diffusion data for arsenic, boron, and phosphorous in silicon gives the expression (Masters and Gorey, 1978)

\[ c_V^{eq} = 5 \times 10^{25} \exp\left(-\frac{3.66 \text{ eV}}{kT}\right) \text{ cm}^{-3}. \] (6.104)

This expression gives predictions for \( c_V^{eq} \) which are slightly lower than the predictions of eq. (6.102). However, the two results essentially overlap when the experimental error is taken into account. Our simulation results have the advantage of being direct within the limitations posed by the SW atomic model compared to the analyses of dopant diffusion data.

The results of our simulations for the self-interstitial are compared with results from the analysis of gold diffusion data by Morehead (1988). The Arrhenius expression for the equilibrium concentration of self-interstitials derived by Morehead is

\[ c_i^{eq} = 1 \times 10^{27} \exp\left(-\frac{3.8 \text{ eV}}{kT}\right) \text{ cm}^{-3}, \] (6.105)

over the temperature range \( 973 \text{ K} < T < 1523 \text{ K} \). Equation (6.105) was the appropriate fit for \( c_i^{eq} \) incorporated in the expression for the effective diffusivity of gold in silicon to reproduce gold concentration profiles by numerical solution of the diffusion equation. Fitting our simulation results for the temperature dependence of \( c_i^{eq} \) by a single Arrhenius expression gives

\[ c_i^{eq} = 1.2 \times 10^{25} \exp\left(-\frac{3.66 \text{ eV}}{kT}\right) \text{ cm}^{-3}. \] (6.106)

However, the number of the internal degrees-of-freedom for the extended interstitial was kept equal to one in our calculations, i.e. \( n_{i,d.o.f.} = 1 \) in eq. (6.84). Therefore, the value of \( 1.2 \times 10^{25} \text{ cm}^{-3} \) in eq. (6.106) corresponds to the appropriate lower bound for the preexponential factor. For this reason and because of the closeness of the activation energies in eqs. (6.105) and (6.106) our simulation results are in reasonable agreement with the analysis of Morehead. The multiplicity of the extended interstitial configurations is discussed in Section 6.3.5.
6.3.5. Diffusion Coefficients and Contributions to Self-Diffusion

Calculation of migration paths for point defect diffusion consists of minimizing the energy of the system under the constraint that some degrees of freedom are eliminated by fixing some of the coordinates of the migrating atom along the path that connects two nearest-neighbor equilibrium lattice sites. Therefore, several energy minimization calculations are required, each corresponding to a specific location that the migrating atom occupies in the lattice space between the original site and the site which is occupied after the jump occurs. These detailed calculations are complementary in the investigation of the diffusion process and clarify mechanisms of migration. The temperature dependence of the migration properties and the diffusion coefficients can be calculated from the results of our simulations for the temperature dependence of the formation properties and application of jump rate theory.

In our investigation of the migration paths all three coordinates of the jumping atom are held fixed during energy minimization, i.e. the jumping particle is restricted to occupy specific points on the line that connects the two equilibrium sites (Shoemaker et al., 1991). Another approach where the jumping particle is restricted on the plane perpendicular to the line that connects the two equilibrium sites also has been followed (Viñals and Sekerka, 1988) reducing the number of degrees of freedom only by one. We follow the first approach in order to couple this study with the previous studies of the strictly localized point defect configurations. In addition, application of both approaches at the midpoint of the vacancy migration path gives the same result for the energy of the relaxed atomic configuration which is the formation energy of the split-vacancy.

The migration path coordinate $\eta$ is defined (Shoemaker et al., 1991) as

$$r = r_1 + \eta (r_2 - r_1),$$

(6.107)

where $r$ is the position vector of the jumping atom and $r_1$ and $r_2$ are the position vectors of the initial and the final equilibrium lattice sites in the jump, respectively. The values of $\eta$ fall in the interval $[0,1]$. Therefore, energy minimization is carried out under the constraint of constant $\eta$. In the case where all the three coordinates of the jumping particle are held
fixed the effective frequency $\tilde{\nu}$ defined in eq. (6.86) is rewritten as a geometric mean taking into account the elimination of three normal modes instead of one:

$$\tilde{\nu} = \left[ \prod_{i=1}^{3N} \frac{\nu_i}{\prod_{i=1}^{3N-3} \nu_i^*} \right]^{1/3}.$$

(6.108)

The unrelaxed configuration for the jump of an atom from an equilibrium site to a nearest neighbor vacant site is shown in Fig. 6.27 together with the results of our simulations for the variation in the local energy minima as a function of the migration path coordinate $\eta$. Two sets of results are shown for the local energy minima. The upper curve corresponds to energy minimization using the steepest descent quench for a system with the volume of the perfect crystal at $p = 0$ and using an unrelaxed diamond lattice as an initial configuration where the location of the jumping atom with respect to the vacant site is given by eq. (6.107). The lower curve corresponds to energy minimization using the steepest descent quench after relaxation using a $(NpT)$ MC simulation at $p = 0$ and $T = 500$ K, i.e. our standard minimization routine. In both cases the energy is referenced to the energy of the perfect crystal.

The results of the two minimization schemes are totally different. The first supports the split-vacancy ($\eta = 0.5$) as the stable vacancy configuration and gives an energy value for the vacancy ($\eta = 0, \eta = 1$) equal almost to the unrelaxed value of 4.34 eV reported by Batra et al. (1987). Our calculation following the first minimization scheme predicts several metastable configurations for different values of $\eta$ with energies above the upper curve of Fig. 6.27. The results of our calculation following the second minimization scheme which has been consistently followed in the calculation of formation energies proves that the split-vacancy is indeed the saddle point configuration of vacancy migration. Minimization using the steepest descent quench alone was motivated by the results of Ghaisas (1991) who claimed that the SW atomic model favors the split-vacancy configuration. Ghaisas also suggested a modification of the SW interatomic potential by strengthening the three-body contributions by a factor of 1.75, i.e. the value of the parameter $\lambda$ of the SW potential in Table 6.1 was multiplied by 1.75. Our calculations of the vacancy migration path prove that such a modification in the SW interatomic potential resulting in questionable changes
Figure 6.27. Atomic jump toward a vacant lattice site. The unrelaxed atomic configuration shows the migration path. The variation of the energy as a function of the migration path coordinate $\eta$ is shown following two different energy minimization techniques.
of other predicted thermophysical properties is not necessary. However, the results of the lower curve of Fig. 6.27 show some metastable configurations where the jumping atom is located close to the equilibrium sites \((\eta < 0.2, \eta > 0.8)\).

The temperature dependence is required of the enthalpy and the Gibbs free energy of migration for the calculation of the vacancy diffusivity according to eqs. (6.91). The migration properties are defined as the differences in the formation properties of the point defect at the equilibrium and the saddle point position. The temperature dependence of the diffusivity is computed using eq. (6.91b). The closeness in the predictions of eqs. (6.91a) and (6.91b) is tested at the reference temperature \(T_{\text{ref}} = 500\ \text{K}\). The effective jump frequency is computed using eq. (6.108) where the extended vacancy is taken as the equilibrium configuration and the split-vacancy as the saddle point configuration. The frequency products in the right side of eq. (6.108) are proportional to the determinants of the dynamical matrices evaluated at the corresponding configurations. The simulation result \(\tilde{\nu} = 1.0539 \times 10^{13}\ \text{s}^{-1}\) is in excellent agreement with the value of \(kT/h = 1.0418 \times 10^{13}\ \text{s}^{-1}\) at \(T = 500\ \text{K}\). Moreover, the closeness in the values of 0.48 eV and 0.46 eV of the predicted enthalpy and Gibbs free energy of migration validates the use of eq. (6.91b) based on the high temperature approximation for the entropy.

Equation (6.91b) using the vacancy as the equilibrium position and the split-vacancy as the saddle point position in the jump is used to calculate the temperature dependence of the vacancy diffusion coefficient \(D_V\). The result is presented in Fig. 6.28 for 250 K \(\leq T \leq 1635\ \text{K}\); for \(T > 1000\ \text{K}\) the vacancy diffusivities are of the order of \(10^{-5}\) to \(10^{-3}\ \text{cm}^2/\text{s}\). These very high values are explained by the low Gibbs free energies of migration. The migration energy of 0.43 eV for the vacancy computed from our simulations as the difference between the zero temperature formation energies of the vacancy and the split-vacancy is in reasonable agreement with the experimental value of 0.33 eV for the migration enthalpy of the neutral vacancy over the temperature range 150 K \(< T < 180\ \text{K}\). Furthermore, the simulation result is in agreement with the experimental value of 0.45 eV for the positively charged vacancy \((V^{2+})\) at \(T = 220\ \text{K}\).

The contributions to self-diffusion from the vacancy mechanism is given by the product \(c_V^q D_V\). The results of the simulations for the vacancy contribution to self-diffusion are

331
Figure 6.28. Temperature dependence of the vacancy diffusion coefficient.
presented in Fig. 6.29. The simulation results are compared with results from analyses of gold diffusion data (Tan and Gösele, 1985; Morehead, 1988). Tan and Gösele (1985) derived the correlation

\[ c_V^q D_V = 0.6 \exp \left( -\frac{4.06 \text{ eV}}{kT} \right) \text{ cm}^2/\text{s}, \quad (6.109) \]

over the temperature range 973 K < T < 1373 K. Morehead derived the correlation

\[ c_V^q D_V = 40 \exp \left( -\frac{4.6 \text{ eV}}{kT} \right) \text{ cm}^2/\text{s}, \quad (6.110) \]

over the temperature range 973 K < T < 1523 K. These results have been extrapolated over the temperature range 900 K < T < 1635 K in Fig. 6.29. The results from the simulations predict a much higher vacancy contribution to self-diffusion compared to the correlations of eqs. (6.109) and (6.110). This is because of the 1 to 1.5 eV lower activation energy predicted by simulation than the activation energies in eqs. (6.109) and (6.110); the activation energy predicted by simulation is equal to the Gibbs free energy of formation for the split-vacancy.

Several migration paths, both simple and interstitially, have been investigated for self-interstitial migration by the total energy calculations of Car et al. (1984). The ab initio total energy calculations of Bar-Yam and Joannopoulos (1984a, b) suggested the low electron density path (TH) connecting two nearest neighbor tetrahedral sites through the hexagonal site as the appropriate path avoiding unstable configurations. Interstitial migration through the low electron density path was originally proposed by Pantelides et al. (1983). Although the results of our calculations support the extended self-interstitial the simple interstitial path TH is investigated in detail. The results for the variation of the energy are presented in Fig. 6.30, referenced to the energy of the perfect crystal, as a function of the migration path coordinate \( \eta \). The predicted migration energy is 1.74 eV. These calculations provide a clear picture for the stability of the tetrahedral interstitial relative to the hexagonal interstitial within the SW framework. The stability of the interstitial configurations according to the ab initio calculations depend on the charge state and the Fermi level (Car et al., 1984, 1985; Bar-Yam and Joannopoulos, 1984a, b).
Figure 6.29. Temperature dependence of the component of the self-diffusion coefficient due to the vacancy mechanism. Comparison between simulation results and analyses of gold diffusion data.
Figure 6.30. Energy variation as a function of the migration path coordinate $\eta$ for the simple interstitial path TH.
The motivation for the detailed investigation of the simple interstitial path TH is the determination of the lower bound for the self-interstitial diffusivity. This is suggested by the high migration energy of 1.74 eV, which is the highest predicted among all the possible simple or interstitialcy paths. Moreover, experimental data (Griffin et al., 1985) have provided lower bounds for interstitial migration which are compared with our simulation results.

A recent estimate of the self-interstitial diffusion coefficient in silicon was given by Wijjanakula (1990) from experimentally measured concentration profiles of oxygen thermal donors for 733 K < T < 773 K. The analysis assumed the validity of the model of Ourmazd et al. (1984) according to which the donor activity arises from the two broken bonds of a silicon interstitial located at the core of the donor and is terminated by emitting the charged silicon interstitial (I²⁺). The diffusion profiles were analyzed assuming that the diffusing species is the self-interstitial. This experimental data at oxygen-donor-formation temperatures were combined with the results of Griffin et al. (1985) from oxidation-enhanced and retarded diffusion in n-type silicon for T > 1223 K. The resulting Arrhenius expression for the interstitial diffusivity is

\[
D_I = 3.35 \times 10^{-1} \exp\left(-\frac{1.86 \text{ eV}}{kT}\right) \text{ cm}^2/\text{s}, \tag{6.111}
\]

over the temperature range 733 K < T < 1473 K. The results of Griffin et al. (1985) are only a lower bound on interstitial migration rates, so eq. (6.111) also is a lower bound for the self-interstitial diffusivity in this temperature range.

The total energy calculations of Car et al. (1984) and Bar-Yam and Joannopoulos (1984a, b) proved that the tetrahedral site is the stable site for I²⁺ in n-type silicon and that migration occurs through the TH path. These results therefore support our interpretation of the experimental data by analysis of interstitial migration along the TH path. Our simulation results for the temperature dependence of the Gibbs free energies of formation of the tetrahedral and the hexagonal self-interstitial are combined with eq. (6.91b) to calculate the temperature dependence of the interstitial diffusivity for the TH path. The jump distance \(d_0\) depends on the lattice parameter \(a\) which is calculated from the temperature dependence of the density of the perfect crystal from our simulations.
The calculated values of the diffusivity of \(3.2 \times 10^{-14}\) and \(1.26 \times 10^{-13}\) cm\(^2\)/s at \(T = 733\) K and \(T = 773\) K, are in reasonable agreement with the corresponding experimental values of \(1.0 \times 10^{-13}\) and \(1.3 \times 10^{-13}\) cm\(^2\)/s (Wijaranakula, 1990). Furthermore, the calculated value of \(2 \times 10^{-8}\) cm\(^2\)/s at \(T = 1473\) K is in good agreement with the value of \(3 \times 10^{-8}\) cm\(^2\)/s reported by Griffin et al. (1985) for the same temperature. The corresponding Arrhenius plot is presented in Fig. 6.31 and shows the good agreement between the result of our simulations and correlation (6.111) over \(673\) K < \(T\) < \(1635\) K. The calculated average value of the entropy of migration over this temperature range is 2.7 \(k\).

The contributions of self-interstitials to self-diffusion is calculated based on the results for the energetically favored extended self-interstitial. In the simulations extended interstitials are formed starting from several initial configurations where the extra atom is put in various sites between the tetrahedral and the hexagonal site. The resulting final quenched configurations all have the formation energy of 3.66 eV and the extra atom occupies several positions that surround the tetrahedral site. These calculations suggest a mechanism for the migration of the extended interstitial. The extra atom moves from one position that corresponds to the low energy extended interstitial to another energetically equivalent position while passing through the lowest energy configuration of high symmetry, i.e. the tetrahedral interstitial site.

This model gives an activation energy of 4.84 eV for the interstitial contribution \(c_1^{9} D_1\) to the self-diffusion coefficient, which is equal to the formation energy of the tetrahedral self-interstitial. This activation energy has been estimated indirectly by analysis of gold diffusion data; The values of \(Q\) reported by Tan and Gösele (1985) and by Morehead (1988) are 4.84 and 5.0 eV, respectively, and are in very good agreement with the model of extended interstitial migration. Moreover, the activation enthalpy for interstitial diffusion estimated by Morehead (1988) was 1.2 eV, also in excellent agreement with the value of 1.18 eV which is the calculated difference in the formation enthalpies of the extended and the tetrahedral interstitial, respectively, at \(p = 0\).

One point that should be clarified in the analysis of the interstitial contribution to self-diffusion is that simple interstitial paths, such as the TH path, do not contribute to substitutional diffusion because they do not affect the migration of atoms that occupy
Figure 6.31. Temperature dependence of the self-interstitial diffusivity for the simple interstitial path TH. Comparison of simulation results with experimental data.
lattice sites. Therefore, only the interstitialcy paths will provide contributions to $D_1$ that should be taken into account in the analysis of self-diffusion and substitutional dopant diffusion. The interstitial migration mechanism that involves the extended interstitial constitutes a generalized interstitialcy path where the pronounced lattice relaxation around the extra atom is equivalent to a generalized mechanism of ejection of lattice atoms into channels of interstitial migration. The significant atomic relaxation around the extended defect defines a region in the crystal where substitutional and interstitial atoms become essentially indistinguishable.

The assumption that $d_0$ is equal to one bond distance is introduced in eq. (6.91b) and the simulation results for the temperature dependence of the Gibbs free energies of formation of the extended and the tetrahedral self-interstitial are combined to calculate the diffusion coefficient $D_1$. This result is combined with the calculated value for $C^{eq}_1$ to give the simulated prediction for the interstitial contribution to self-diffusion. The results are presented in the Arrhenius plot of Fig. 6.32 over the temperature range $900 \text{ K} < T < 1635 \text{ K}$. The results are compared with the correlations reported by Tan and Gösele (1985)

$$c^{eq}_1 D_1 = 914 \exp\left(-\frac{4.84 \text{ eV}}{kT}\right) \text{ cm}^2/\text{s}, \quad (6.113)$$

and the correlation reported by Morehead (1988)

$$c^{eq}_1 D_1 = 4 \times 10^3 \exp\left(-\frac{5.0 \text{ eV}}{kT}\right) \text{ cm}^2/\text{s}, \quad (6.114)$$

over the temperature range $973 \text{ K} < T < 1523 \text{ K}$. Although the simulation curve is roughly parallel to the straight lines from correlations (6.113) and (6.114) the simulation predictions are lower by several orders of magnitude. The reason is partly that the observed multiplicity of the extended interstitial configuration is not introduced in the calculation of $c^{eq}_1$.

The favorable interstitialcy path predicted by the SW interatomic potential is the TS path with a migration energy of 0.54 eV. Although this path gives a reasonable value of 5.38 eV for the activation energy of the contribution to self-diffusion $c^{eq}_1 D_1$, the equilibrium concentration of the tetrahedral site is very low due to the high formation energy of

339
Figure 6.32. Temperature dependence of the component of the self-diffusion coefficient due to the interstitialcy mechanism. Comparison between simulation results and analyses of gold diffusion data.
4.84 eV, and the lower configurational entropy caused by the only one internal degree-of-freedom. Therefore, it leads to a much lower interstitial contribution to self-diffusion.

6.4. Summary and Discussion

The role of intrinsic point defects in silicon in determining the behavior of microelectronic devices has been well recognized and these defects have been studied in detail via experimental measurements at low and high temperatures, total energy ab initio calculations, numerical analyses of diffusion data, and lattice statics simulations. The contribution of this thesis to this research is a systematic analysis of the structural, thermodynamic, and transport properties of intrinsic point defects approach based on atomistic simulations using the Stillinger-Weber interatomic potential. The computational method provides a robust approach for calculation of point defect properties over the temperature range from absolute zero to the melting point based on a combination of lattice statics simulations, Monte Carlo simulations, the quasi-harmonic approximation, and the cumulant analysis of the enthalpy distribution function. The main advantage of this approach is that it allows for direct and internally-consistent calculations of the temperature dependence of thermal equilibrium concentrations and diffusion coefficients of vacancies and self-interstitials in silicon.

The results of our calculations for self-interstitial properties are in good agreement with experimental data and clarify mechanisms of interstitial diffusion. The stable interstitial configuration is an extended self-interstitial with a formation energy of about 3.7 eV, which is in agreement with the best available estimate of interstitial equilibrium concentrations (Morehead, 1988). A migration mechanism for the extended interstitial through a saddle point configuration which involves the tetrahedral site gives a migration energy of 1.2 eV again in excellent agreement with analyses of gold diffusion data (Morehead, 1988). Combining the two results gives an activation energy of about 5 eV in very good agreement with the activation energy of the interstitial contribution to self-diffusion in all the correlations that have appeared in the literature (Tan and Gösele, 1985; Fahey et al., 1989). Lengthy molecular dynamics simulations using the Stillinger-Weber poten-
tial will provide further evidence for the migration mechanism of the extended interstitial. A systematic analysis of extended interstitial configurations by \textit{ab initio} calculations or atomistic simulations has never appeared. Finally, calculation of the diffusion coefficient for the TH interstitial path reproduced with good accuracy interstitial diffusion coefficients over a broad temperature range where effects such as thermal donor formation or diffusion influenced by surface oxidation are significant. For the calculation jump rate theory is combined with atomistic calculations of free energies and a detailed analysis of the migration path.

The results of our calculations for vacancy properties suggest that vacancies dominate diffusion processes. The calculated values of the vacancy contribution to self-diffusion are higher than the predictions from analyses of gold diffusion data. This is due to the very low formation energy of about 2.7 eV that is calculated in the simulations. However, the simulation predictions for the thermal equilibrium concentrations of vacancies are intermediate between results from different analyses of positron annihilation experiments (Dannefaer et al., 1986). At high temperatures and taking the experimental error into account the simulation results are in reasonable agreement with the experimental data. In addition, the calculated migration enthalpy of the vacancy of 0.43 eV is in reasonable agreement with the experimental value of 0.33 eV for the neutral vacancy. Furthermore, detailed calculation of the vacancy migration path shows that the split-vacancy is indeed the saddle point configuration and that the Stillinger-Weber interatomic potential does not need to be modified in order to reproduce vacancy migration properties; this result is counter to the claims of Ghaisas (1991).

A final point of discussion involves the phase diagram at $p = 0$. This is presented in Fig. 6.33 and shows the temperature dependence of the density of the perfect crystal, a crystal with a vacancy concentration $x_v = 1/216$, two crystals with interstitial concentrations $x_I = 1/216$ and $x_I = 1/64$, respectively, and the amorphous structure of Wooten, Wainer and Wearie (1985). The similarities in the densities of the amorphous system and the crystal with the higher interstitial concentration shown in Fig. 6.33 are accompanied by similarities in the coordination number and the vibrational entropies (Broughton and Li, 1987). Given that the interstitials in these crystals are extended configurations with
respect to the perfect crystal motivates studies for the transition from the crystalline to the amorphous state via self-interstitials in silicon under supersaturation using the Stillinger-Weber interatomic potential. Further discussion for this mechanism of amorphization cannot be justified without additional simulation results.

Figure 6.33. Phase diagram of silicon at zero pressure from MC simulations using the SW interatomic potential including the perfect crystal, the amorphous phase, and crystals that contain vacancies and self-interstitials at different concentrations.
7. CONCLUSIONS

The focus of this thesis has been a theoretical analysis of the dynamics of intrinsic point defects, impurities, and dislocations in single semiconductor crystals produced by bulk melt crystal growth methods. The studies are based on a systematic combination of atomistic simulations, constitutive modeling, and continuum analysis. This integrated approach establishes the links between the mechanisms of formation and the microscopic dynamics of crystalline defects, as well as the interactions between various types of defects and the macroscopic response of the material under the thermomechanical conditions of processing. The development of constitutive theories that relate micromechanical properties to the stress field and the temperature history of the crystal is the necessary step to link our understanding of the microscopic structure and dynamics of lattice imperfections and impurities with well established classical continuum mechanical formalisms, such as ideal plasticity theory. This theoretical framework guarantees microscopic interpretation of the underlying thermomechanical processes and quantitative prediction of the macroscopic behavior of materials.

In spite of the tremendous experimental and theoretical advances in the field of defect dynamics in semiconductor crystals over the past four decades a quantitative understanding of the role of processing conditions during crystal growth and device processing on the type and density of lattice defects has been lacking. Moreover, the coupled dynamics of point and line defects, the multiplicity of length and time scales associated with the mechanisms of defect formation, propagation and multiplication, and the limitations in computer power make the quantitative and accurate modeling of defect dynamics extremely difficult. The strategy that has been followed in this thesis is based on determination of the dominant defect and the corresponding rate controlling or limiting step during a process and the detailed analysis of the microscopic, mesoscopic, and macroscopic aspects of the relationship between the nature of the defect and the mechanisms of the process. Therefore, the problem of dislocation multiplication during bulk melt growth has been modeled by focusing on a thin crystalline region next to the melt/crystal interface where the growing crystal is considered as undoped, which decouples the interactions between dopants and
dislocations from the analysis of dislocation propagation and multiplication. Impurity-
dislocation interaction is studied for a specific binary system of oxygen in silicon in the
dilute limit of impurity concentration and at uniform temperature field. The quantitative
modeling of the effects of oxygen on the dynamics of an isolated dislocation is the first step
in the development of a constitutive theory of solution hardening of a lattice characterized
by an active dislocation field, which controls the mechanical behavior of the crystal. Finally,
targeted atomistic simulations of the formation and migration of isolated intrinsic point
defects in silicon are carried out to determine the crucial parameters that control self-
diffusion and substitutional dopant diffusion.

The modeling of dislocation dynamics in this thesis goes beyond the simplified ap-
proach of computing the thermal elastic stress field and focuses on the microscopic rate
processes that control crystal plasticity and the establishment of the length and time scales
that characterize such processes. The analysis of impurity-dislocation interaction is based
on a theory of dislocation drag, which takes into account both impurity atmospheres and
impurity clusters inside the dislocation core; these effects lead to hysteresis in the velocity-
stress relation. Simulation of plastic deformation is preceded by constitutive modeling
of solution hardening and dislocation multiplication. Finally, investigation of the struc-
tural, thermodynamic, and transport properties of intrinsic point defects are based on an
integrated atomistic approach that allows for predictions of diffusivities and equilibrium
concentrations of vacancies and self-interstitials at the high temperatures that are of in-
terest in device processing. The conclusions of this thesis are summarized in Section 7.1
and directions for future research are given in Section 7.2.

7.1. Summary

Dislocation dynamics in a growing crystal is analyzed using the model of Alexander
and Haasen (1968) for plastic deformation in diamond/zincblende structure semiconduc-
tors. The thermal stress fields used in the analysis are indicative of the liquid encapsulated
Czochralski (LEC) growth of InP and GaAs and of the Czochralski (CZ) growth of silicon.
For typical crystal growth conditions, the rapid rate of dislocation multiplication compared
to the rate of motion of the crystal through the thermal stress field leads to asymptotic expressions for the bulk dislocation density that depend on the variation of the stress field in the direction of crystal growth. The asymptotic results agree quantitatively with computations and experimental results for dislocation densities in LEC grown InP (Völkl and Müller, 1989) and give simple criteria for reducing dislocation densities in as-grown compound semiconductor crystals.

The asymptotic limit in our analysis is expressed by two dimensionless groups: the Dislocation Damköhler Number $Da \gg 1$ which expresses the rate of dislocation multiplication relative to the rate of crystal motion, and the dimensionless back stress $\delta \ll 1$ which scales the back stress due to dislocation interactions with the elastic stress. The high values of $Da$ for typical growth conditions lead to the formation of a boundary layer in the dislocation density next to the melt/crystal interface. The sharp increase of the dislocation density inside the boundary layer leads to high dislocation densities in the bulk. If the von Mises stress $\sigma$ decreases with the distance from the melt/crystal interface, the bulk dislocation density is proportional to the square of the von Mises stress adjacent to the melt/crystal interface. This is typically the case for the Czochralski growth of silicon and the LEC growth of InP and GaAs with shallow encapsulant layers. If the von Mises stress increases with the distance from the interface the bulk dislocation density scales as $\delta \sigma^2$. Therefore, the dislocation density is again proportional to the square of the von Mises stress, but the very low proportionality factor $\delta \ll 1$ guarantees lower levels of dislocation densities in the bulk. This case is typical of the LEC growth of III-V compound semiconductors with deep encapsulant layers and gives a simple criterion for the design of crystal growth systems for producing crystals with low dislocation densities avoiding the introduction of dopants in the solid phase.

The interaction between impurities and dislocations is studied for the case of interstitial oxygen in CZ grown silicon. A systematic theory is developed for the analysis of oxygen migration near a 60° dislocation and for the resulting retardation of dislocation motion. Quantitative predictions are based on the solution of the macroscopic equation for oxygen transport in the elastic stress field created by the dislocation. The oxygen concentration field is computed numerically using the Galerkin finite-element method. The link
between the microscopic dynamics of interstitial oxygen within the diamond lattice and macroscopic transport is established by a constitutive model for the dependence of the drift velocity and diffusivity of oxygen on the elastic interaction between the dislocation and the solute and on temperature. Constitutive modeling is based on the moment analysis of the transport equation and computation of the first and second moments of the number density distribution function from dynamic Monte Carlo simulations. Drag on the dislocation motion is modeled using the oxygen distribution around the dislocation core, the elastic interaction forces and the energy needed to dissolve oxygen clusters that are formed within the core. The model predicts dislocation locking due to oxygen. The locking and unlocking of dislocations is explained by a hysteresis that is predicted in the dislocation velocity as a function of applied stress. The predictions of the model agree quantitatively with experimental data for dislocation mobility in CZ silicon (Imai and Sumino, 1983).

Constitutive models also are developed that describe the effects of oxygen on the internal stresses in the crystal and on the multiplication of dislocations due to the cross-slip mechanism. Combining the results of the dislocation drag model with the constitutive models for lattice hardening and dislocation multiplication reproduces the experimental stress-strain curves (Yonenaga et al., 1984) and explains the initial stage of deformation of oxygen-hardened silicon. Increasing the concentration of dissolved interstitial oxygen in silicon leads to significant increase of the upper yield stress of silicon. The implications in defect engineering during device processing are important: increased oxygen concentration and dissolution of oxygen precipitates by preannealing at high temperatures prevents the warpage of wafers during thermal cycling.

Thermodynamic, structural and transport properties of intrinsic point defects, i.e. vacancies and self-interstitials, in crystalline silicon are predicted using an integrated approach based on atomistic simulations. Crystalline silicon is modeled by the Stillinger-Weber interatomic potential (Stillinger and Weber, 1985). Lattice statics, lattice dynamics, and Monte Carlo (MC) simulations are combined to calculate crucial point defect parameters for defect engineering. Thermal equilibrium concentrations of point defects are calculated from Gibbs free energy computations based on a combination of the quasiharmonic approximation and the analysis of the first few cumulants of the enthalpy dis-
tribution function computed from isothermal-isobaric MC simulations. The temperature
dependence of formation enthalpies and volumes of point defects are also computed using
the same approach. The zero-temperature structure and energetics of point defects also
are studied using a combination of the simulated annealing and the steepest descent meth-
ods for energy minimization. Important point defect parameters are calculated, such as
formation energies and activation energies of migration. Diffusion coefficients of vacancies
and self-interstitials are computed based on jump rate theory and the simulation results
for enthalpies and Gibbs free energies of migration.

Migration energies and thermal equilibrium concentrations of vacancies are in rea-
sonable agreement with experimental data. Formation and migration properties of self-
interstitials are in good agreement with experimental data. The calculations reproduce
experimentally measured interstitial diffusivities and explain diffusion mechanisms of self-
interstitials and the contributions of self-interstitials to self-diffusion and substitu-
tional dopant diffusion. The development of a robust computational framework for the atomic
modeling of intrinsic point defects at low and high temperatures and the ability of the
interatomic potential of Stillinger and Weber for satisfactory predictions of point defect
properties suggests the application of the integrated approach for atomistic simulations
followed in this research for the investigation of more complicated problems associated
with point defects in silicon.

7.2. Directions for Future Research

Specific topics for further research have arisen during the course of the investigation of
the dynamics of intrinsic point defects, impurities, and dislocations in single semiconductor
crystals grown from the melt. The emphasis in this discussion of future research topics
is put on the important problem of the formation of large-scale defects and disordered
structures in silicon from intrinsic point defects with concentrations that exceed the thermal
equilibrium concentrations.
The Role of Interstitials in the Crystalline – to – Amorphous Transition in Silicon. The similarities between amorphous silicon and crystalline silicon with high concentrations of self-interstitials were discussed in Section 6.4. The results of our calculations for the temperature dependence of the density, the vibrational entropy and the energies at low temperatures of systems with high concentrations of extended self-interstitials suggest a study of the transition from crystalline to amorphous structure via self-interstitials. The role of self-interstitials in solid-state amorphization is a question of general interest (Shoemaker et al., 1991). This is a problem of technical interest, especially in silicon, considering the evidence of injection of self-interstitials as a result of surface oxidation (Tan and Gösele, 1985, Fahey et al., 1989). The similarities between amorphous silicon and crystalline silicon supersaturated in self-interstitials were observed in this study from investigations of small periodically repeated supercells that contain single self-interstitials. The absence of periodicity in the amorphous phase suggests a study of structural, energetic and vibrational properties of systems represented by large simulation cells that contain thousands of hosts atoms and a large number of extra atoms which are randomly distributed in interstitial sites. Using large system sizes also permits the increase in the number of interstitial atoms in the system while keeping the interactions between interstitial atoms at a minimum level to avoid clustering. Calculation of the density at zero temperature, the energy at low and high temperatures, and global and local structure factors is a straightforward task within the computational framework developed in the present thesis. However, increase of the system size leads to tremendous memory requirements in the construction and the analysis of the dynamical matrix for calculation of vibrational properties. Therefore, the efficient use of the sparsity of the dynamical matrix \( \mathbf{D} \) is a major requirement; the form of \( \mathbf{D} \) is discussed in Section A.2 of the Appendix.

Formation of Stable Clusters of Intrinsic Point Defects in Silicon. Lattice statics simulations and free energy calculations are powerful methods for the study of the formation and the stability of clusters of vacancies and self-interstitials in bulk silicon. The binding energies of these clusters are easily calculated by energy minimization of the system that contains the cluster of point defects and the system that contains an equal number of point defects randomly distributed in the bulk. An efficient technique for the investigation of
cluster stability is the minimization of the Helmholtz free energy within the local harmonic approximation (LeSar et al., 1989). This technique avoids the storage problems associated with the dynamical matrix and combines system relaxation with free energy calculation of the relaxed structure. The success of this technique for the study of extended defects such as grain boundaries and the distribution of solute on surfaces and grain boundaries has been demonstrated by Wang et al. (1991) and provides the motivation for the use of the local harmonic approximation for clusters of point defects in spite of the associated anharmonicities. Monitoring the binding energy of the cluster and the free energy difference between the system that contains the cluster and the system that contains the random distribution of point defects as a function of the cluster size will provide the information for the onset of cluster stability and the size of the critical nucleus for cluster growth provided that the system is sufficiently supersaturated in point defects. In addition, application of homogeneous nucleation theory (Turnbull and Fisher, 1949; Woodruff, 1973) provides the basis for estimation of the growth rate of stable clusters. However, if the concentration of vacancies or self-interstitials is kept constant at some set level of supersaturation the calculations require the continuous increase in the size of the simulation cell and the computations become very intensive.

The Onset of Dislocation Formation. Calculation of the onset of dislocation formation in silicon follows the same lines with the study of formation of stable clusters of intrinsic point defects. An additional element of major importance in this calculation is the structural and geometrical rules that have to be followed in the construction of the cluster of vacancies and self-interstitials. An intuitive structural rule that results from the structure of dislocations and the study of crystallographic glide in diamond-cubic crystals is the removal of neighboring atoms in {111} planes for vacancy clusters and the addition of interstitial atoms in planes parallel to {111} planes. Continuous use of this procedure for a large simulation cell that includes free surfaces will produce an array of parallel macroscopic dislocations by insertion or removal of atomic half-planes. However, after reaching a critical size the clusters of intrinsic point defects are expected to collapse to microscopic dislocation loops surrounded by a stacking fault ribbon. This is expected to be observed as a result of energy minimization of the system that contains the defects. The experience
of analogous lattice statics simulations in metals (Bullough and Perrin, 1968; Sabochick et al., 1988) is valuable for such large-scale computations. The structural characterization of the resulting dislocation loops will provide insights for the mechanisms of microdefect formation in silicon. Available experimental observations of microdefect structures by high resolution transmission electron microscopy (Chikawa, 1987) will serve as guides for the size and structure of these microdefects.

Mesoscopic and Macroscopic Modeling of Point Defects in Bulk Melt Growth Systems.

Improvement of processing conditions has resulted in dislocation free silicon (Dash, 1959). For silicon crystals grown by the Czochralski method the macroscopic modeling of intrinsic point defects is at least of equal importance with dislocation modeling in III-V compound semiconductors. This study will relate the results of the previous atomistic calculations of point defects and clusters of point defects with realistic processing conditions such as growth velocities and temperature gradients. The outcome of this study will be the distribution of intrinsic point defects and microdefects in the growing silicon crystal. A first step toward this direction in point defect and microdefect modeling in silicon crystal growth was made by Voronkov (1982) based on phenomenological models. The necessary database for the macroscopic model will be provided by the atomistic simulations of intrinsic point defects and clusters of intrinsic point defects. Two of the necessary components for this macroscopic model - equilibrium concentrations and diffusivities of vacancies and self-interstitials - are provided by the results in this thesis. The analysis of the growth of point defect clusters will provide data for the evaluation of rates of point defect production and destruction, while the knowledge of the thermal equilibrium concentrations as functions of temperature with respect to the actual concentrations of intrinsic point defects provide the measure of the driving force for cluster formation. Two elements are still required for the development of a realistic macroscopic model for point defects and microdefects during growth. The first is the knowledge of the recombination rate between vacancies and self-interstitials. Evaluation of this rate requires lengthy molecular dynamics simulations of Frenkel pair formation and destruction in a system that contains both vacancies and self-interstitials. Again experience from atomistic simulations in metals will be valuable in such computations (Kwok et al., 1984). These simulations also will provide excellent tests for
the diffusion mechanisms that are suggested in this thesis from the combination of results of static and equilibrium simulations with jump-rate theory. The second element is the knowledge of the intrinsic point defect concentrations at the melt/crystal interfacial region. The assumption of Voronkov (1982) that these point defects are in thermal equilibrium at the melting temperature has been examined in Section 2.4. However, the difficulty of atomistic simulation in the melt/crystal interfacial region during growth at very low growth velocities makes rigorous atomistic calculations impractical. The development of a model analogous to the work of Bolling and Fainstein (1972) is suggested for the interfacial region. The parameters of this model, eq. (2.11), can be taken from atomistic simulations for the bulk liquid phase (Broughton and Li, 1987) and for the structure of the melt/crystal interface (Landmar. et al., 1986).

In conclusion, the suggested macroscopic modeling of point defects and microdefects will retain the key features of the macroscopic modeling of dislocations which are the division of the crystal to interfacial and bulk region and the evaluation of the coefficients of transport and kinetic equations from microscopic studies. The main difference is that, instead of the use of phenomenological constants that characterizes the model of Alexander and Haasen for the analysis of dislocation dynamics, thermodynamic and transport properties of point defects and clusters of point defects will be used. These properties are evaluated from targeted atomistic simulations using the interatomic potential of Stillinger and Weber (1985).
REFERENCES


Reichl L. E., 'Modern Course in Statistical Physics', University of Texas Press, Austin (1980).


A.1. Interatomic Forces and Internal Virial

Calculation of the interatomic forces is required for energy minimization using the steepest descent method. In this calculation the descent direction $-\mathbf{g}$ of eq. (6.39) is given by a $3N$-dimensional vector, where $N$ is the number of particles in the simulation cell and $\mathbf{g}$ is the gradient of the potential energy $\Phi$ of the system. The gradient is defined as

$$
\mathbf{g} \equiv \nabla \Phi = -[f_1, f_2, \ldots, f_N]^T,
$$

(A.1)

where $f_n, n = 1, 2, \ldots, N$ denotes the force exerted on atom $n$ from its interaction with all the other atoms in the simulation cell. The interatomic distances are calculated following the minimum image convention (Allen and Tildesley, 1989).

Using an equivalent form of eq. (6.15) for the potential energy $\Phi$ and taking the gradient with respect to the coordinates of particle $n$ gives the three-dimensional force components $\mathbf{f}_n$ of the vector in the right-hand side of (A.1) as

$$
\mathbf{f}_n = -\nabla_n \Phi = -\frac{1}{2} \sum_i \sum_{j \neq i} \nabla_n f_2(r_{ij}) - \frac{1}{6} \sum_i \sum_{j \neq i} \sum_{k \neq j \neq i} \nabla_n f_3(r_i, r_j, r_k),
$$

(A.2)

where the Stillinger-Weber interatomic potential functions are given by equations (6.12) and (6.13) for the two-body and the three-body interactions, respectively. Energies and lengths in equation (A.2) are dimensionless using the energy scale $\epsilon$ and the length scale $\sigma$ of the interatomic potential, respectively. The gradient of the potential of the two-body interaction $f_2$ with respect to the coordinates of particle $n$ is

$$
\nabla_n f_2(r_{ij}) = \frac{\partial f_2}{\partial \mathbf{r}_i} \delta_{ni} + \frac{\partial f_2}{\partial \mathbf{r}_j} \delta_{nj} = \frac{\partial f_2}{\partial r_{ij}} \hat{r}_{ij} \delta_{ni} - \frac{\partial f_2}{\partial r_{ij}} \hat{r}_{ij} \delta_{nj},
$$

(A.3)

where $\delta$ is the Kronecker delta, and $\hat{r}_{ij}$ is the unit vector in the direction $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$. Newton’s third law has been used to switch indices in the right-hand side of (A.3). Therefore, calculation of the two-body forces reduces to the calculation of the derivative of $f_2(r)$ with respect to $r$.

375
Equation (6.13) for the three-body interaction is rewritten as

\[ f_3(r_i, r_j, r_k) = h_1 + h_2 + h_3, \quad (A.4) \]

where the functions \( h_1, h_2 \) and \( h_3 \) have exactly the same functional form given by eq. (6.13b) but different arguments as indicated by eq. (6.13a). Therefore, the gradient of \( f_3 \) with respect to the coordinates of particle \( n \) is written as

\[ \nabla_n f_3 = \frac{\partial h_1}{\partial r_i} \delta_{ni} + \frac{\partial h_1}{\partial r_j} \delta_{nj} + \frac{\partial h_1}{\partial r_k} \delta_{nk} + \frac{\partial h_2}{\partial r_i} \delta_{ni} + \frac{\partial h_2}{\partial r_j} \delta_{nj} + \frac{\partial h_2}{\partial r_k} \delta_{nk} + \frac{\partial h_3}{\partial r_i} \delta_{ni} + \frac{\partial h_3}{\partial r_j} \delta_{nj} + \frac{\partial h_3}{\partial r_k} \delta_{nk}. \quad (A.5) \]

Equation (6.13b) gives the gradients of \( h_1 \) with respect to the coordinates of atoms \( i, j \) and \( k \), respectively, as

\[ \frac{\partial h_1}{\partial r_i} = \left[ \frac{\partial h_1}{\partial r_{ij}} + \left( \frac{1}{r_{ik}} - \frac{\mu_i}{r_{ij}} \right) \frac{\partial h_1}{\partial \mu_i} \right] \dot{r}_{ij} + \left[ \frac{\partial h_1}{\partial r_{ik}} + \left( \frac{1}{r_{ij}} - \frac{\mu_i}{r_{ik}} \right) \frac{\partial h_1}{\partial \mu_i} \right] \dot{r}_{ik}, \quad (A.6a) \]

\[ \frac{\partial h_1}{\partial r_j} = - \left[ \frac{\partial h_1}{\partial r_{ij}} + \left( \frac{1}{r_{ij}} - \frac{\mu_i}{r_{ij}} \right) \frac{\partial h_1}{\partial \mu_i} \right] \dot{r}_{ij} - \frac{r_{jk}}{r_{ij} r_{ik}} \frac{\partial h_1}{\partial \mu_i} \dot{r}_{jk}, \quad (A.6b) \]

\[ \frac{\partial h_1}{\partial r_k} = - \left[ \frac{\partial h_1}{\partial r_{ik}} + \left( \frac{1}{r_{ij}} - \frac{\mu_i}{r_{ik}} \right) \frac{\partial h_1}{\partial \mu_i} \right] \dot{r}_{ij} + \frac{r_{jk}}{r_{ij} r_{ik}} \frac{\partial h_1}{\partial \mu_i} \dot{r}_{jk}, \quad (A.6c) \]

where

\[ \mu_i \equiv \cos \theta_{ijk} = \frac{r_{ij} \cdot r_{ik}}{r_{ij} r_{ik}} = \frac{x_{ij} x_{ik} + y_{ij} y_{ik} + z_{ij} z_{ik}}{r_{ij} r_{ik}}. \quad (A.6d) \]

Equations (A.6) satisfy Newton's third law which is expressed here by

\[ \frac{\partial h_1}{\partial r_i} + \frac{\partial h_1}{\partial r_j} + \frac{\partial h_1}{\partial r_k} = 0. \quad (A.7) \]

The gradients of the components \( h_2 \) and \( h_3 \) with respect to the coordinates of atoms \( i, j \) and \( k \) are given by equations which are derived from (A.6) by circular interchange of the indices \( i, j, \) and \( k \).

The calculation of the internal virial \( \mathcal{W} \) of eq. (6.36) is carried out by splitting the interaction forces to two-body and three-body contributions as

\[ \mathcal{W} \equiv \frac{1}{3} \sum_{i=1}^{N} r_i \cdot f_i = \frac{1}{3} \sum_{i=1}^{N} r_i \cdot f_{2,i} + \frac{1}{3} \sum_{i=1}^{N} r_i \cdot f_{3,i} = \mathcal{W}_2 + \mathcal{W}_3. \quad (A.8) \]
Following the analysis of Allen and Tildesley (1989) gives for the contribution of the two-body interaction to the internal virial

$$\mathcal{W}_2 = -\frac{1}{3} \sum_i \sum_{j>i} r_{ij} \frac{\partial f_2}{\partial r_{ij}}.$$  (A.9)

Following a similar analysis for the contribution of the three-body interaction to the internal virial and using eqs. (A.6) for the three-body forces gives

$$\mathcal{W}_3 = -\frac{1}{3} \sum_i \sum_{j>i} \sum_{k>j>i} r_{ij} \left( \frac{\partial h_1}{\partial r_{ij}} + \frac{\partial h_2}{\partial r_{ij}} \right) + r_{ik} \left( \frac{\partial h_1}{\partial r_{ik}} + \frac{\partial h_3}{\partial r_{ik}} \right) + r_{jk} \left( \frac{\partial h_2}{\partial r_{jk}} + \frac{\partial h_3}{\partial r_{jk}} \right).$$  (A.10)

Equations (6.34) and (A.8)-(A.10) are used to monitor the variation in the pressure of the system during Monte Carlo simulations in the (NVT) or (NpT) ensembles.

A.2. Calculation of the Dynamical Matrix

The calculation of the dynamical matrix, i.e. the Hessian matrix of the potential energy evaluated at the equilibrium configuration is carried out according to equations (6.65), (6.71), and (6.72). The FORTRAN subroutine “calcimat” presented in this Section is used for the computation of the elements $D_{ij,\alpha\beta}$ of the dynamical matrix.
subroutine calcmat(tote,cube,cube2,np)

This subroutine calculates the elements of the dynamical matrix D.

The two-body contributions and the three-body contributions are calculated separately.
The off-diagonal elements are calculated first in both cases according to equation (6.71).
Then the diagonal part of the matrix is calculated according to equation (6.72).
This diagonal part corresponds to the local (3X3) dynamical matrices of all particles.
This subroutine also calculates the total energy of the atomic configuration (tote) using as input the length of the edge of the cubic simulation cell (cube) and the number of atoms in the cell (np).

cube2 = cube / 2

dyn(i,j)=element of dynamical matrix

x0(i),y0(i),z0(i): coordinates of particle 'i'

The parameters of the SW potential are given by the common blocks phi2par and phi3par

The function xygimg is used for the calculation of the minimum image interatomic distances

This version of the subroutine is not optimized.
The Verlet neighbor list is not used
The three-body loops do not make full use of the symmetry of the potential energy functions

Dynamical Matrix Calculation

implicit double precision (a-h,o-z)
parameter npd=301
common/dmat/dyn(3*npd,3*npd)
common/pos/x0(npd),y0(npd),z0(npd)
common/propa/np1,np2,np22,rmax,bolk,eps
common/phi2par/acap,bcap,p,q,a
common/phi3par/al,gamma

xyzimg(w,wr)=(wr-cube)*dsign(1.d0,w)

tote=0.d0
tote2=0.d0
tote3=0.d0

xi=0.d0
yi=0.d0
zi=0.d0
xj=0.d0
yj=0.d0
zj=0.d0
xk=0.d0
yk=0.d0
zik=0.d0

indi=0
inds=0

Outer loop over atoms
do 450 i=1,np
imat=3*(i-1)+1

    c.....Store position of atom i
    xi=x0(i)
yi=y0(i)
zi=z0(i)

    c.....First inner loop over atoms
    do 440 j=1,np
       if (j.eq.i) goto 440
       jmat=3*(j-1)+1
       f2r=0.0d0
       xij=xj-x0(j)
yij=yj-y0(j)
zij=zi-z0(j)

       c.....Find image of atom j closest to atom i
       xrij=dabs(xij)
yrij=dabs(yij)
zrij=dabs(zij)
       if (xrij.gt.cube2) xij=xzim(xij,xrij)
       if (yrij.gt.cube2) yij=xzim(yij,yrij)
       if (zrij.gt.cube2) zij=xzim(zij,zrij)

       c.....Calculate r(ij)
       rijsq=xij**2+yij**2+zij**2
       rijct=dsqrt(rijsq)

    c.....Store position of atom j
    xj=x0(j)
yj=y0(j)
zj=z0(j)
       if (rijsq.gt.rmax) then
          indij=1
          goto 440
       else
          indij=0
       endif

    c.....Calculations for dynamical matrix elements
    c.....Step #1
    c.....2-body Contributions
    r=rijct

    c.....Calculation of energy, first and second order derivatives
    f2r=acap*(bcap/r**p-1.0d0)*dexp(1.0d0/(r-a))
tote2=tote2+f2r
    df2dr=-f2r/(r-a)**2-acap*p*bcap*dexp(1.0d0/(r-a))/(r**2-(p+1)
    d2f2dr=2.0d0*f2r/(r-a)**3-df2dr/(r-a)**2 +
           acap*p*bcap*dexp(1.0d0/(r-a))/((r-a)**2*r**2/(p+1)) +
           acap*bcap*p*(p+1)*dexp(1.0d0/(r-a))/r**2/(p+2)

    c.....Analytical derivatives of interatomic distance wrt
    c.....atomic coordinates to be used in the calculation of
    c.....matrix elements
    drdxi=xij/r
drdxj=-xij/r
drdyi=yij/r
drdyj=-yij/r
drdzj=zi/j/r
drdzj=-zij/r
d2rdxidxj=((xij/r)**2-1.0d0)/r
d2rdxidyj=xij*yij/r**3
d2rdxidzj=xij*zi/j/r**3
d2rdyidyj=xij*yij/r**3
d2rdyidyj=((yij/r)**2-1.0d0)/r
d2rdyidyj=zi/j/r**3

d2rdzidxj=xij*zij/r**3
d2rdzidyj=yij*zij/r**3
d2rdzidzj=((zij/r)**2-1.0d0)/r

d2rdxidxi=(1.0d0-(xij/r)**2)/r

d2rdyidyj=(1.0d0-(yij/r)**2)/r

d2rdzidzi=(1.0d0-(zij/r)**2)/r

d2rdxidyj=xij*yij/r**3
d2rdxidzi=xij*zij/r**3
d2rdyidzi=yij*zij/r**3

c c.....Two-body contributions to off-diagonal elements

c dyn (imat, jmat)=d2f2dr2*drdxi*drdxi+jdf2dr*d2rdxidxj
  dyn (imat, jmat+1)=d2f2dr2*drdxi*drdij+df2dr*d2rdxidyj
  dyn (imat, jmat+2)=d2f2dr2*drdxi*drdzj+df2dr*d2rdxidzj
  dyn (imat+1, jmat)=d2f2dr2*drdyj*drdxi+df2dr*d2rdyidyj
  dyn (imat+1, jmat+1)=d2f2dr2*drdyj*drdij+df2dr*d2rdyidyj
  dyn (imat+1, jmat+2)=d2f2dr2*drdyj*drdzj+df2dr*d2rdyidzj
  dyn (imat+2, jmat)=d2f2dr2*drdzj*drdxi+df2dr*d2rdzidxj
  dyn (imat+2, jmat+1)=d2f2dr2*drdzj*drdij+df2dr*d2rdzidyj
  dyn (imat+2, jmat+2)=d2f2dr2*drdzj*drdzj+df2dr*d2rdzidzj

c c.....Diagonal elements (2-body)

c dyn (imat, imat)=dyn (imat, imat)+d2f2dr2*drdxi*drdxi+
+ df2dr*d2rdxidxi
  dyn (imat, imat+1)=dyn (imat, imat+1)+d2f2dr2*drdxi*drdij+
+ df2dr*d2rdxidyj
  dyn (imat, imat+2)=dyn (imat, imat+2)+d2f2dr2*drdxi*drdzj+
+ df2dr*d2rdxidzj
  dyn (imat+1, imat)=dyn (imat+1, imat)+d2f2dr2*drdij*drdxi+
+ df2dr*d2rdyidyj
  dyn (imat+1, imat+1)=dyn (imat+1, imat+1)+d2f2dr2*drdij*drdij+
+ df2dr*d2rdyidyj
  dyn (imat+1, imat+2)=dyn (imat+1, imat+2)+d2f2dr2*drdij*drdzj+
+ df2dr*d2rdyidzj
  dyn (imat+2, imat)=dyn (imat+2, imat)+d2f2dr2*drdzj*drdxi+
+ df2dr*d2rdzidxj
  dyn (imat+2, imat+1)=dyn (imat+2, imat+1)+d2f2dr2*drdzj*drdij+
+ df2dr*d2rdzidyj
  dyn (imat+2, imat+2)=dyn (imat+2, imat+2)+d2f2dr2*drdzj*drdzj+
+ df2dr*d2rdzidzj
  indi=indi+1
440  continue
450  continue
  print *, 'Energy of 2-body interactions: E/N=', tote2/(2*np)

c c.....3-body Contributions

c
do 500 i=1,np
  imat=3*(i-1)+1

c.....Store position of atom i
  xi=x0(i)
  yi=y0(i)
  zi=z0(i)
do 490  j=1,np
  if (j.eq.i) goto 490
  jmat=3*(j-1)+1
  xij=xi-x0(j)
  yij=yi-y0(j)
  zij=zi-z0(j)

c.....Find image of atom j closest to atom i
  xrij=dabs(xij)
  yrij=dabs(yij)
zrij=dabs(zij)
if (xrij.gt.cube2) xij=xyzimg(xij,xrij) 
if (yrij.gt.cube2) yij=xyzimg(yij,yrij) 
if (zrij.gt.cube2) zij=xyzimg(zij,zrij)

c.....Calculate r(ij)
rijsq=xij**2+yij**2+zij**2
rijctr=dsqrt(rijsq)

c.....Store position of atom j
xj=x0(j)
yj=y0(j)
zj=z0(j)
if (rijsq.gt.rmax) then
   indij=1
   goto 460
else
   indij=0
endif

c
460

  do 480 k=1,np
    if (k.eq.i.or.k.eq.j) goto 480
  c

    hi0=0.d0
    hjk=0.d0
    hki=0.d0
    f3xixi=0.d0
    f3xiyi=0.d0
    f3xiyi=0.d0
    f3xiyj=0.d0
    f3xiyj=0.d0
    f3xizj=0.d0
    f3yiij=0.d0
    f3yizi=0.d0
    f3yixj=0.d0
    f3yixj=0.d0
    f3yizj=0.d0
    f3zizi=0.d0
    f3zizj=0.d0
    f3zijj=0.d0
    f3zijj=0.d0
    f3rijjk=0.d0
  c

c.....Distance between atom j and atom k
xjk=xj-x0(k)
yjk=yj-y0(k)
zjk=zj-z0(k)

c.....Find image of atom k closest to atom j
xrkj=dabs(xjk)
yrkj=dabs(yjk)
zrkj=dabs(zjk)
if (xrkj.gt.cube2) xjk=xyzimg(xjk,xrkj)
if (yrkj.gt.cube2) yjk=xyzimg(yjk,yrkj)
if (zrkj.gt.cube2) zjk=xyzimg(zjk,zrkj)

c.....Calculate r(jk)
riksq=xjk**2+yjk**2+zjk**2
rkjctr=dsqrt(riksq)
if (riksq.gt.rmax) then
   indjk=1
   goto 470
else
   indjk=0
   if (indij.eq.1) goto 470
endif

c.....Calculate angle thijk
if (indij.eq.0.and.indjk.eq.0) then
dnumijk=-xij*xjk-yij*yjk-zij*zjk
domegaijk=rijctr*rkjctr
Calculate 3-body interaction given by potential function \( h_2 \).

Calculate first and second derivatives

Derivatives of interatomic distances wrt particle coordinates

\[
\begin{align*}
rl &= r_{1jk} \cr
r3 &= r_{jk3} \cr
th2 &= \cos \theta_{ijk} \cr
drl dx_i &= x_j/r_l \cr
drl dx_j &= x_i/r_l \cr
drl dy_i &= y_j/r_l \cr
drl dy_j &= y_i/r_l \cr
drl dz_i &= z_j/r_l \cr
drl dz_j &= z_i/r_l \cr

d2rl dx_idxj &= (x_j/r_l)^2 - 1.0 \cr
d2rl dx_idyj &= x_j y_j/r_l^3 \cr
d2rl dx_idzj &= x_j z_j/r_l^3 \cr
d2rl dy_idxj &= y_j x_j/r_l^3 \cr
d2rl dy_idyj &= y_j y_j/r_l^3 \cr
d2rl dy_idzj &= y_j z_j/r_l^3 \cr
d2rl dz_idxj &= x_j z_j/r_l^3 \cr
d2rl dz_idyj &= y_j z_j/r_l^3 \cr
d2rl dz_idzj &= (z_j/r_l)^2 - 1.0 \cr
d2rl dxidxi &= (1.0 - (x_j/r_l)^2) \cr
d2rl dyidi &= (1.0 - (y_j/r_l)^2) \cr
d2rl dzidzi &= (1.0 - (z_j/r_l)^2) \cr
d2rl dxidyi &= -x_j y_j/r_l^3 \cr
d2rl dzidyi &= -x_j z_j/r_l^3 \cr
d2rl dyidi &= -y_j z_j/r_l^3 \cr
d3dx_j &= x_{jk}/r_3 \cr
d3dy_j &= y_{jk}/r_3 \cr
d3dz_j &= z_{jk}/r_3
\end{align*}
\]

Derivatives of 3-body angle cosines wrt particle coordinates

First-order derivatives

\[
\begin{align*}
dth2 dx_i &= x_j/(r_l r_3) - th2*x_j/r_l^2 \cr
dth2 dx_j &= (x_j - x_i)/(r_l r_3) - th2*(x_j/r_l^2) \cr
dth2 dy_i &= y_j/(r_l r_3) - th2*y_j/r_l^2 \cr
dth2 dy_j &= (y_j - y_i)/(r_l r_3) - th2*(y_j/r_l^2) \cr
dth2 dz_i &= z_j/(r_l r_3) - th2*z_j/r_l^2 \cr
dth2 dz_j &= (z_j - z_i)/(r_l r_3) - th2*(z_j/r_l^2)
\end{align*}
\]

Second-order derivatives

\[
\begin{align*}
dth2 dx_i x_i &= -dth2 dx_i*(x_j/r_l^2) - th2*(r_l^2 - 2.0*x_j/r_l^2)/(r_l r_3) \cr
dth2 dx_i y_i &= dth2 dx_i*(y_j/r_l^2) + th2* \cr
dth2 dx_i z_i &= dth2 dx_i*(z_j/r_l^2) + th2* \cr
dth2 dx_i x_j &= -dth2 dx_j*(x_i/r_l^2) + th2* \cr
dth2 dx_i y_j &= dth2 dx_j*(y_i/r_l^2) + th2* \cr
dth2 dx_i z_j &= dth2 dx_j*(z_i/r_l^2) + th2*
\end{align*}
\]

\[
\begin{align*}
dth2 dx_i x_j &= +x_j y_j/(r_l^2) \cr
dth2 dx_i y_j &= +x_j z_j/(r_l^2) \cr
dth2 dx_i z_j &= +x_j z_j/(r_l^2) \cr
dth2 dy_i x_i &= -dth2 dy_i*(x_j/r_l^2) - th2* \cr
dth2 dy_i y_i &= dth2 dy_i*(y_j/r_l^2) + th2* \cr
dth2 dy_i z_i &= dth2 dy_i*(z_j/r_l^2) + th2*
\end{align*}
\]

\[
\begin{align*}
dth2 dy_i x_j &= +x_j y_j/(r_l^2) \cr
dth2 dy_i y_j &= +x_j z_j/(r_l^2) \cr
dth2 dy_i z_j &= +x_j z_j/(r_l^2)
\end{align*}
\]
Energy contribution

\[ \text{hjk} = a l \cdot \text{dexp}(\gamma/(r_1-a) + \gamma/(r_3-a)) \cdot (\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot 2 \]
\[ \text{f}_{3rijk} = f_{3rijk} + \text{hjk} \]

Contributions to the gradient (first derivatives)

\[ \text{fmuljk} = a l \cdot \text{dexp}(\gamma/(r_1-a) + \gamma/(r_3-a)) \cdot (\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot 2 \]
\[ \text{hjkxi} = \text{fmuljk} \cdot (-2.0 \cdot d_0 \cdot yj/(r_1 \cdot r_3) - 2.0 \cdot d_0 \cdot \text{th}_2 \cdot xj/(r_1 \cdot r_3)^2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot xj/(r_1 - (r_1 - a) \cdot 2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot yj/(r_3 - (r_3 - a) \cdot 2) \]
\[ \text{hjkk} = (-\text{hjkxi} + \text{hjkk}) \]
\[ \text{hjky} = \text{fmuljk} \cdot (-2.0 \cdot yj/(r_1 \cdot r_3) - 2.0 \cdot d_0 \cdot \text{th}_2 \cdot yj/(r_1 \cdot r_3)^2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot yj/(r_1 - (r_1 - a) \cdot 2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot yj/(r_3 - (r_3 - a) \cdot 2) \]
\[ \text{hjky} = (-\text{hjky} + \text{hjky}) \]
\[ \text{hkzi} = \text{fmuljk} \cdot (-2.0 \cdot zj/(r_1 \cdot r_3) - 2.0 \cdot d_0 \cdot \text{th}_2 \cdot zj/(r_1 \cdot r_3)^2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot zj/(r_1 - (r_1 - a) \cdot 2) \]
\[ + \gamma/(\text{th}_2 + 1.0 \cdot d_0/3.0) \cdot zj/(r_3 - (r_3 - a) \cdot 2) \]
\[ \text{hjkz} = (-\text{hkzi} + \text{hjkz}) \]

Contributions to the Hessian (second derivatives)

\[ \text{hjkxixi} = \text{hjkxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dxi}/(r_1-a)^2) + \]
\[ \gamma/(\text{dr}_1 \cdot \text{dxi}/(r_1-a)^3) \]
\[ + \gamma/(\text{dr}_1 \cdot \text{dxi}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dxi}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dxi} \cdot \text{dth}_2 \cdot \text{dxi} + 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dxi} \cdot \text{dxi} \]
\[ \text{hjkxiyi} = \text{hjkxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dyi}/(r_1-a)^2) + \]
\[ \gamma/(\text{dr}_1 \cdot \text{dyi}/(r_1-a)^3) \]
\[ + \gamma/(\text{dr}_1 \cdot \text{dyi}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dyi}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyi} \cdot \text{dth}_2 \cdot \text{dyi} + 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyi} \cdot \text{dyi} \]
\[ \text{hjkxizj} = \text{hjkxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^3) \]
\[ + \gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot \text{dth}_2 \cdot \text{dzj} + 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot \text{dzj} \]
\[ \text{hjkxjyj} = \text{hjkxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dyj}/(r_1-a)^2) + \gamma/(\text{dr}_1 \cdot \text{dyj}/(r_1-a)^3) \]
\[ + \gamma/(\text{dr}_1 \cdot \text{dyj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyj} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dyj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyj} \cdot \text{dth}_2 \cdot \text{dyj} + 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dyj} \cdot \text{dyj} \]
\[ \text{hjkxjzj} = \text{hjkxi} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^3) \]
\[ + \gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot (-\gamma/(\text{dr}_1 \cdot \text{dzj}/(r_1-a)^2) + \]
\[ 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot \text{dth}_2 \cdot \text{dzj} + 2.0 \cdot d_0 \cdot \text{fmuljk} \cdot \text{dth}_2 \cdot \text{dzj} \cdot \text{dzj} \]
a)**2)+hjk*(2.0*gamma*drldzi*/(r-l)**2-
+ gamma*d2rldxldzj*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyiy=hjkyi*(-gamma*drldzi*/(r-l)**2)+
+ hjk*(2.0*gamma*drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyizj=hjkyi*(-gamma*drldzi*/(r-l)**2)+
+ hjk*(2.0*gamma*drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyixj=hjkyi*(-gamma*drldzi*/(r-l)**2-
+ 2-gamma*dr3dxi*/(r-l)**2)+
+ hjk*(2.0*gamma*_drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyizj=hjkyi*(-gamma*drldzi*/(r-l)**2-
+ 2-gamma*dr3dxi*/(r-l)**2)+
+ hjk*(2.0*gamma*_drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyizj=hjkyi*(-gamma*drldzi*/(r-l)**2-
+ 2-gamma*dr3dxi*/(r-l)**2)+
+ hjk*(2.0*gamma*_drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyixj=hjkyi*(-gamma*drldzi*/(r-l)**2-
+ 2-gamma*dr3dxi*/(r-l)**2)+
+ hjk*(2.0*gamma*_drldzi*/(r-l)**2-
+ gamma*d2rldydyi*/(r-l)**2)+
+ 2.0*fmuljk*dth2dxi*(-gamma*drldzi*/(r-l)**2)+
+ 2.0*fmul2*dth2dxi*dth2dxi+2.0*fmuljk*dth2dxi
hjkyiy=f3xiyi
f3xii=j3xii+yjxj
f3xiji=f3xiji+yjxj
f3yiyi=f3yiyi+yjyj
f3zizi=f3zizi+yjzi
f3ziji=f3ziji+yjzi
f3ziyj=f3ziyj+yjzi
f3ziji=f3ziji+yjzi
f3ziyj=f3ziyj+yjzi
f3ziji=f3ziji+yjzi
c.....Distance between atom i and atom k
470  xik=x1-x0(k)
yik=y1-y0(k)
zik=z1-z0(k)
c.....Find image of atom k closest to atom i
xrik=dabs(xik)
yrik=dabs(yik)
zrik=dabs(zik)
if (xrik.gt.cube2) xik=xyzimg(xik,xrik)
if (yrik.gt.cube2) yik=xyzimg(yik,yrik)
if (zrik.gt.cube2) zik=xyzimg(zik,zrik)
c.....Calculate r(ik)
riksq=xik**2+yik**2+zik**2
rikctr=dsqrt(riksq)
if (riksq.gt.rmax) then
  indik=1
  goto 479
else
  indik=0
  if (indij.eq.1.and.indjk.eq.1) goto 479
endif

c.....Calculate angle thjik
if (indij.eq.0) then
dnumjik=xij*xik+yij*yik+zij*zik
denomjik=rijctr*rikctr
cosjik=dnumjik/denomjik
thi=cosjik
c

c.....Calculate 3-body interaction given by potential function h1.
c.....Calculate first and second derivatives
c
c..... Derivatives of interatomic distances wrt particle coordinates
c
  r1=rijctr
  r2=rikctr
  drlldxi=xij/r1
  drlldxj=-xij/r1
  drlldyi=yij/r1
  drlldyj=-yij/r1
  dr'dzi=zij/r1
  drlldzj=-zij/r1
  d2rlldxj=(xij/r1)**2-1.0d0)/r1
  d2rlldyj=xij*yij/r1**3
  d2rlldzj=xij*zij/r1**3
  d2rlldyj=(yij/r1)**2-1.0d0)/r1
  d2rlldyj=yij*zij/r1**3
  d2rlldyj=xij*zij/r1**3
  d2rlldyj=yij*zij/r1**3
  d2rlldyj=((zij/r1)**2-1.0d0)/r1
  d2rlldyj=1.0d0-(xij/r1)**2)/r1
  d2rlldyj=1.0d0-(yij/r1)**2)/r1
  d2rlldzj=1.0d0-(zij/r1)**2)/r1
  d2rlldyj=-xij*yij/r1**3
  d2rlldyj=-xij*zij/r1**3
  d2rlldyj=-yij*zij/r1**3
  dr2dxi=xik/r2
  dr2dyi=yik/r2
  dr2dzi=zik/r2
  d2r2dxidxi=1.0d0-(xik/r2)**2)/r2
  d2r2dyidyi=1.0d0-(yik/r2)**2)/r2
  d2r2dzidzi=1.0d0-(zik/r2)**2)/r2
  d2r2dxidyi=-(xik*yik)/r2**3
  d2r2dzidzi=-(xik*zik)/r2**3
  d2r2dyidyi=-(yik*zik)/r2**3
Derivatives of 3-body angle cosines wrt particle coordinates

First-order derivatives

\[
dhldx_i = \frac{(x_i + x_2)}{(r_1 r_2)} - \theta_1 \left( \frac{(x_i / r_1^2 + x_2 / r_2^2)}{r_1^2} \right) + \theta_1 \left( \frac{x_2 / r_1^2}{r_1^2} \right) - \left( \frac{x_2 / r_1^2}{r_1^2} \right)
\]

Second-order derivatives

\[
dhldx_i = \frac{2.0 (x_i / r_1^2)}{(r_1 r_2)} - dhldx_i \left( \frac{x_i / r_1^2 + x_2 / r_2^2}{r_1^2} \right) + \left( \frac{2.0 (x_i / r_1^2)}{r_1^2} \right) - \left( \frac{x_2 / r_1^2}{r_1^2} \right) - \left( \frac{x_2 / r_1^2}{r_1^2} \right)
\]

Energy contribution

\[
h_{ij} = a \cdot dexp(\gamma / (r_1 - a) + \gamma / (r_2 - a)) \cdot (\theta_1 + 1.0 / 3.0) \cdot (r_1^2 + r_2^2)
\]

Contributions to the gradient (first derivatives)

\[
f_{ij} = a \cdot dexp(\gamma / (r_1 - a) + \gamma / (r_2 - a)) \cdot (\theta_1 + 1.0 / 3.0)
\]

\[
h_{ij} = \frac{f_{ij}}{(r_1 - a) + \gamma / (r_2 - a)}
\]
hijx = fmulij(-2.0*d0*xij/(r1*r2)+2.0*d0*th1*xik/(r2**2)
+ gamma*(th1+1.0*d0/3.0)*xik/(r2*(r2-a)**2))

hijxi = (hijx+hijxk)

hijy = fmulij(-2.0*d0*yij/(r1*r2)+2.0*d0*th1*yij/(r2**2)
+ gamma*(th1+1.0*d0/3.0)*yij/(r2*(r2-a)**2))

hijy = (hijy+hijyk)

hijz = fmulij(-2.0*d0*zij/(r1*r2)+2.0*d0*th1*zij/(r2**2)
+ gamma*(th1+1.0*d0/3.0)*zij/(r2*(r2-a)**2))

hijz = (hijz+hijzk)

hijx = fmulij(-gamma*dr1dx/(r1-a)**2-
gamma*dr2dxi/(r2-a)**2)

hij = (2.0*d0*gamma*dr1dx*dr1dx/(r1-a)**3-
gamma*dr2dxidxi/(r1-a)**2+
2.0*d0*gamma*dr2dxidxi/(r2-a)**3-
gamma*dr22dxiidxi/(r2-a)**2)

hijxi = (hijx+hijxk)

hijy = fmulij(-gamma*dr1dy/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)

hij = (2.0*d0*gamma*dr1dy*dr1dy/(r1-a)**3-
gamma*dr2dxidyi/(r1-a)**2+
2.0*d0*gamma*dr2dxidyi/(r2-a)**3-
gamma*dr22dxidyi/(r2-a)**2)

hijyi = (hijy+hijyk)

hijz = fmulij(-gamma*dr1dz/(r1-a)**2-
gamma*dr2dxi/(r2-a)**2)

hij = (2.0*d0*gamma*dr1dz*dr1dz/(r1-a)**3-
gamma*dr2dxidzi/(r1-a)**2+
2.0*d0*gamma*dr2dxidzi/(r2-a)**3-
gamma*dr22dxidzi/(r2-a)**2)

hijzi = (hijz+hijzk)

hijxj = fmulij(-gamma*dr1dxj/(r1-a)**2+
gamma*dr2dxijdj/(r1-a)**2)

hij = (2.0*d0*gamma*dr1dxj*dr1dxj/(r1-a)**3-
gamma*dr2dxijdijdj/(r1-a)**2+
2.0*d0*gamma*dr2dxijdijdj/(r2-a)**3-
gamma*dr22dxijdijdj/(r2-a)**2)

hijxj = (hijxj+hijxjz)

hijyj = fmulij(-gamma*dr1dyj/(r1-a)**2+
gamma*dr2dyijdj/(r1-a)**2)

hij = (2.0*d0*gamma*dr1dyj*dr1dyj/(r1-a)**3-
gamma*dr2dyijdijdj/(r1-a)**2+
2.0*d0*gamma*dr2dyijdijdj/(r2-a)**3-
gamma*dr22dyijdijdj/(r2-a)**2)

hijyj = (hijyj+hijyjz)

hijzj = fmulij(-gamma*dr1dzj/(r1-a)**2+
gamma*dr2dxijdjzj/(r1-a)**2)

hij = (2.0*d0*gamma*dr1dzj*dr1dzj/(r1-a)**3-
gamma*dr2dxjidzjzj/(r1-a)**2+
2.0*d0*gamma*dr2dxjidzjzj/(r2-a)**3-
gamma*dr22dxjidzjzj/(r2-a)**2)

hijzj = (hijzj+hijzjx)
2.0*d0*gamma*dr2dyi*dr2dyi/(r2-a)**3-
gamma*dr2dyidyi/(r2-a)**2+
2.0*d0*fmulij*dthldyi*(-gamma*drldyi/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)+
2.0*d0*fmulij*dthldyi*dthldyi+2.0*fmulij*dthldyi
hiyizi=hiyi*(-gamma*drldzi/(r1-a)**2-
gamma*dr2dzi/(r2-a)**2)+
hij*(2.0*d0*gamma*drldyi*drldzi/(r1-a)**3-
gamma*dr2ldydzi/(r1-a)**2)+
2.0*d0*gamma*dr2dyi*dr2dzi/(r2-a)**3-
gamma*dr2dyidzi/(r2-a)**2)+
2.0*fmulij*dthldzi*(-gamma*drldyi/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)+
2.0*fmulij*dthldyi*dthldzi+2.0*fmulij*dthldyi
hiyijx=hiyi*(-gamma*drldxj/(r1-a)**2)+
hij*(2.0*d0*gamma*drldyi*drldxj/(r1-a)**3-
gamma*dr2ldydxdj/(r1-a)**2)+
2.0*fmulij*dthldxj*(-gamma*drldyi/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)+
2.0*fmulij*dthldyi*dthldxj+2.0*fmulij*dthldyi
hiyizj=hiyi*(-gamma*drldzj/(r1-a)**2)+
hij*(2.0*d0*gamma*drldyi*drldzj/(r1-a)**3-
gamma*dr2ldydzj/(r1-a)**2)+
2.0*fmulij*dthldzj*(-gamma*drldyi/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)+
2.0*fmulij*dthldyi*dthldzj+2.0*fmulij*dthldyi
hiyziz=hiyi*(-gamma*drldzi/(r1-a)**2-
gamma*dr2dzi/(r2-a)**2)+
hij*(2.0*d0*gamma*drldzi*drldzi/(r1-a)**3-
gamma*dr2ldzdzi/(r1-a)**2)+
2.0*gamma*dr2dzi*dr2dzi/(r2-a)**3-
gamma*dr2zdtdzi/(r2-a)**2)+
2.0*fmulij*dthldzi*(-gamma*drldzi/(r1-a)**2-
gamma*dr2dzi/(r2-a)**2)+
2.0*fmulij*dthldzi*dthldzi+2.0*fmulij*dthldzi
hiyzix=hiyi*(-gamma*drldzi/(r1-a)**2)+
hij*(2.0*d0*gamma*drldzi*drldzi/(r1-a)**3-
gamma*dr2ldzdzi/(r1-a)**2)+
2.0*fmulij*dthldzi*(-gamma*drldzi/(r1-a)**2-
gamma*dr2dzi/(r2-a)**2)+
2.0*fmulij*dthldzi*dthldzi+2.0*fmulij*dthldzi
hiyziy=hiyi*(-gamma*drldyi/(r1-a)**2)+
hij*(2.0*d0*gamma*drldyi*drldyi/(r1-a)**3-
gamma*dr2ldydzi/(r1-a)**2)+
2.0*gamma*dr2dyi*dr2dyi/(r2-a)**3-
gamma*dr2dyidzi/(r2-a)**2)+
2.0*fmulij*dthldyi*(-gamma*drldyi/(r1-a)**2-
gamma*dr2dyi/(r2-a)**2)+
2.0*fmulij*dthldyi*dthldyi+2.0*fmulij*dthldyi
```c
f3yizi=f3yizi+hijyizi
f3yixj=f3yixj+hijyixj
f3yiyj=f3yiyj+hijyiyj
f3yzij=f3yzij+hijyizj
f3zizi=f3zizi+hijzizi
f3zixj=f3zixj+hijzixj
f3ziyj=f3ziyj+hijziyj
f3zizj=f3zizj+hijzizj

endif

.....Calculate angle thikj
if (indjk.eq.0) then
    dnumikj=xik*xjk+yik*yjk+zik*zjk
    denomikj=rikcrt*rjkct
    cosikj=dnumikj/denomikj
    th3=cosikj

.....Calculate 3-body interaction given by potential function h3.
.....Calculate first and second derivatives

.....Derivatives of interatomic distances wrt particle coordinates

r2=rikcrt
r3=rjkct
dr2dxi=xik/r2
dr2dxy=yik/r2
dr2dzui=zik/r2
dr22dxdxi=(1.d0-(xik/r2)**2)/r2
dr22dyidyi=(1.d0-(yik/r2)**2)/r2
dr22dxdzidzi=(1.d0-(zik/r2)**2)/r2
dr22dxidyi=-(xik*yik)/r2**3
dr22dxidzi=-(xik*zik)/r2**3
dr22dyidyzi=-(yik*zik)/r2**3
dr3dxj=xjk/r3
dr3dyj=yjk/r3
dr3dzj=zjk/r3

.....Derivatives of 3-body angle cosines wrt particle coordinates

.....First-order derivatives

dth3dxi=xjk/(r2*r3)-th3*xik/r2**2
dth3dxj=xik/(r2*r3)-th3*xjk/r3**2
dth3dyi=yjk/(r2*r3)-th3*yik/r2**2
dth3dyj=yik/(r2*r3)-th3*yjk/r3**2
dth3dzui=zik/(r2*r3)-th3*zik/r2**2
dth3dzxj=zjk/(r2*r3)-th3*zjk/r3**2

.....Second-order derivatives

dth3dxdxi=-dth3dxi*(xik/r2**2)-th3*(r2**2-
                   2.d0*xik**2)/r2**4-xjk*(xik/r2**2)/(r2*r3)
              + 2.d0*xik**2)/r2**4-xjk*(xik/r2**2)/(r2*r3)
    + dth3dxiyi=-dth3dxiyi*(yik/r2**2)+th3*
              + 2.d0*xik**2)/r2**4-yjk*(xik/r2**2)/(r2*r3)
    + dth3dxiizi=-dth3dxiizi*(zik/r2**2)+th3*
              + 2.d0*xik**2)/r2**4-zjk*(xik/r2**2)/(r2*r3)
    + dth3dxiizj=-xik*(xik/r2**2)/(r2*r3)-dth3dxi*(xjk/r3**2)+
                   1.d0/(r2*r3)
    + dth3dyiyj=-yik*(xik/r2**2)/(r2*r3)-dth3dxiyi*(yjk/r3**2)
    + dth3dxizj=-zik*(xik/r2**2)/(r2*r3)-dth3dxixi*(zik/r3**2)
    + dth3dyizj=-dth3dyizj*(zik/r2**2)+th3*
              + 2.d0*xik**2)/r2**4-zjk*(yik/r2**2)/(r2*r3)
    + dth3dyiyj=-yik*(yik/r2**2)/(r2*r3)+th3*
              + 2.d0*xik**2)/r2**4-yjk*(yik/r2**2)/(r2*r3)
```
\[ 1.0/(r2*r3) \]
\[ \text{dth3dyizj} = \text{zik} \ast (\text{yik}/r2**2)/(r2*r3) - \text{dth3dyj} \ast (zjk/r3**2) \]
\[ \text{dth3d3zi} = \text{zik} \ast (\text{zjk}/r2**2) - \text{th3} \ast (r2**2 - 2.0)*zik**2)/(r2**2) - zjk* (zik/r2**2)/(r2*r3) \]
\[ \text{dth3d3zi} = \text{zik} \ast (\text{yik}/r2**2)/(r2*r3) - \text{dth3d3zi} \ast (yjk/r3**2) \]
\[ \text{dth3d3zi} = \text{zik} \ast (\text{zik}/r2**2)/(r2*r3) - \text{dth3d3zi} \ast (zik/r3**2) + 1.0/(r2*r3) \]

---

**Energy contribution**

\[ \text{hki} = a1 \ast \text{dexp(\gamma)/(r2-a)+\gamma/(r3-a))} \ast (\text{th3}+1.0/3.0) \times 2 \]
\[ \text{f3rij} = f3rij + \text{hki} \]

---

**Contributions to the gradient (first derivatives)**

\[ \text{fmulki} = a1 \ast \text{dexp(\gamma)/(r2-a)+\gamma/(r3-a))} \ast (\text{th3}+1.0/3.0) \]
\[ \text{fmul3} = a1 \ast \text{dexp(\gamma)/(r2-a)+\gamma/(r3-a))} \]
\[ \text{hkixi} = \text{fmulki} \ast (2.0 \times \text{yik}/r2*r3) - 2.0 \ast \text{th3} \times \text{zik} \ast (r2**2) \]
\[ - \text{gamma} \ast (\text{th3}+1.0/3.0) \times \text{zik} \ast (r2**2) \]
\[ + \text{hkixj} = \text{fmulki} \ast (2.0 \times \text{zik} \ast (r2*r3) - 2.0 \ast \text{th3} \ast \text{yjk} \ast (r3**2) \]
\[ - \text{gamma} \ast (\text{th3}+1.0/3.0) \times \text{yjk} \ast (r3**2) \]
\[ + \text{hkiy} = \text{fmulki} \ast (2.0 \times \text{zik}/r2*r3) - 2.0 \ast \text{th3} \times \text{yjk} \ast (r3**2) \]
\[ - \text{gamma} \ast (\text{th3}+1.0/3.0) \times \text{yjk} \ast (r3**2) \]
\[ + \text{hkiy} = \text{fmulki} \ast (2.0 \times \text{zik}/r2*r3) - 2.0 \ast \text{th3} \times \text{yjk} \ast (r3**2) \]
\[ - \text{gamma} \ast (\text{th3}+1.0/3.0) \times \text{yjk} \ast (r3**2) \]
\[ + \text{hkizj} = \text{fmulki} \ast (2.0 \times \text{zik}/r2*r3) - 2.0 \ast \text{th3} \times \text{yjk} \ast (r3**2) \]
\[ - \text{gamma} \ast (\text{th3}+1.0/3.0) \times \text{yjk} \ast (r3**2) \]

---

**Contributions to the Hessian (second derivatives)**

\[ \text{hkixi} = \text{hki} \ast (-\text{gamma} \ast \text{dr2dxi}/(r2-a)**2) + \]
\[ \text{hkixi} \ast (2.0 \ast \text{gamma} \ast \text{dr2dxi} \ast \text{dr2dxi}/(r2-a)**3 - \text{gamma} \ast \text{dr2dxi} \ast \text{dr2dxi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3dxi} \ast (-\text{gamma} \ast \text{dr2dxi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3dxi} \ast \text{dth3dxi} + 2.0 \ast \text{fmulki} \ast \text{dth3dxi} \]
\[ \text{hkixyi} = \text{hkixi} \ast (-\text{gamma} \ast \text{dr2dyi}/(r2-a)**2) + \]
\[ \text{hkixi} \ast (2.0 \times \text{gamma} \ast \text{dr2dx} \ast \text{dr2dyi}/(r2-a)**3 - \text{gamma} \ast \text{dr2dx} \ast \text{dr2dyi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3dyi} \ast (-\text{gamma} \ast \text{dr2dx}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3dyi} \ast \text{dth3dyi} + 2.0 \ast \text{fmulki} \ast \text{dth3dxi} \]
\[ \text{hkixzi} = \text{hkixi} \ast (-\text{gamma} \ast \text{dr2d3zi}/(r2-a)**2) + \]
\[ \text{hkixi} \ast (2.0 \times \text{gamma} \ast \text{dr2d3xi} \ast \text{dr2d3zi}/(r2-a)**3 - \text{gamma} \ast \text{dr2d3xi} \ast \text{dr2d3zi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3d3zi} \ast (-\text{gamma} \ast \text{dr2d3xi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3d3zi} \ast \text{dth3d3zi} + 2.0 \ast \text{fmulki} \ast \text{dth3dxi} \]
\[ \text{hkixj} = \text{hkixi} \ast (-\text{gamma} \ast \text{dr3dxi}/(r3-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3d3xj} \ast (-\text{gamma} \ast \text{dr3d3xj}/(r3-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3d3xj} \ast \text{dth3d3xj} + 2.0 \ast \text{fmulki} \ast \text{dth3d3xj} \]
\[ \text{hkixyj} = \text{hkixi} \ast (-\text{gamma} \ast \text{dr3dxi}/(r3-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3d3yi} \ast (-\text{gamma} \ast \text{dr3d3yi}/(r3-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3d3yi} \ast \text{dth3d3yi} + 2.0 \ast \text{fmulki} \ast \text{dth3d3yi} \]
\[ \text{hkiyj} = \text{hkixi} \ast (-\text{gamma} \ast \text{dr2d3yi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmulki} \ast \text{dth3d3yi} \ast (-\text{gamma} \ast \text{dr2d3yi}/(r2-a)**2) + \]
\[ 2.0 \ast \text{fmul3} \ast \text{dth3d3yi} \ast \text{dth3d3yi} + 2.0 \ast \text{fmulki} \ast \text{dth3d3yi} \]
2.0*fmulki*dth3dyj + hkiyj = hkiyj* (-gamma*dr3dzj/(r3-a)**2) +
2.0*fmulki*dth3dxj* (-gamma*dr2dyi/(r2-a)**2) +
2.0*fmulki*dth3dyj* (-gamma*dr2dyi/(r2-a)**2) +
2.0*fmulki*dth3dyj* (-gamma*dr3dzj/(r3-a)**2) +
2.0*fmulki*dth3dyj* (-gamma*dr3dzj/(r3-a)**2) +
2.0*fmulki*dth3dyi* (-gamma*dr3dzyj + hkiyj = hkiyj* (-gamma*dr3dzj/(r3-a)**2) +
2.0*fmulki*dth3dyi* (-gamma*dr3dzj/(r3-a)**2) +
2.0*fmulki*dth3dyj* (-gamma*dr2dyi/(r2-a)**2) +
2.0*fmulki*dth3dyi* (-gamma*dr2dyi/(r2-a)**2) +
2.0*fmulki*dth3dyi* (-gamma*dr3dzyj +
2.0*fmulki*dth3dyi* (-gamma*dr3dzyj +
2.0*fmulki*dth3dyi* (-gamma*dr3dzyj +
2.0*fmulki*dth3dyi* (-gamma*dr3dzyj +

f3xixi = f3xixi + hkiyj
f3xiyi = f3xiyi + hkiyj
f3xizi = f3xizi + f3xixi
f3xixj = f3xixj + hkiyj
f3xiyj = f3xiyj + hkiyj
f3xizj = f3xizj + hkiyj
f3yiyi = f3yiyi + hkiyj
f3yizi = f3yizi + f3yiyi
f3yixj = f3yixj + hkiyj
f3yiyj = f3yiyj + hkiyj
f3ziyi = f3ziyi + f3zixi
f3ziyj = f3ziyj + hkiyj
f3zixj = f3zixj + hkiyj
f3ziyi = f3ziyi + hkiyj
f3zixj = f3zixj + hkiyj
f3zizi = f3zizi + f3zixi
f3zixj = f3zixj + hkiyj
f3ziyj = f3ziyj + hkiyj
f3zixj = f3zixj + hkiyj
f3zizi = f3zizi + f3zixi

endif

c
479
tote3 = tote3 + f3rijk
inds = inds + 1

c.....Off-diagonal elements (3-body)

dyn (imat, jmat) = dyn (imat, jmat) + f3xixj
dyn (imat, jmat+1) = dyn (imat, jmat+1) + f3xiyj
dyn (imat, jmat+2) = dyn (imat, jmat+2) + f3xizj
dyn (imat+1, jmat) = dyn (imat+1, jmat) + f3yixj
dyn (imat+1, jmat+1) = dyn (imat+1, jmat+1) + f3yiyj
dyn (imat+1, jmat+2) = dyn (imat+1, jmat+2) + f3yizj
dyn (imat+2, jmat) = dyn (imat+2, jmat) + f3zixj
dyn (imat+2, jmat+1) = dyn (imat+2, jmat+1) + f3ziyj
dyn (imat+2, jmat+2) = dyn (imat+2, jmat+2) + f3zizj

c.....Diagonal elements (3-body)

dyn (imat, imat) = dyn (imat, imat) + 0.5d0*f3xixi
dyn (imat, imat+1) = dyn (imat, imat+1) + 0.5d0*f3xiyi
dyn (imat, imat+2) = dyn (imat, imat+2) + 0.5d0*f3xizi
dyn (imat+1, imat) = dyn (imat+1, imat) + 0.5d0*f3xiyi
dyn (imat+1, imat+1) = dyn (imat+1, imat+1) + 0.5d0*f3xiyi
dyn (imat+1, imat+2) = dyn (imat+1, imat+2) + 0.5d0*f3xizi
dyn (imat+2, imat) = dyn (imat+2, imat) + 0.5d0*f3xiyi

dyn (imat+2, imat+1) = dyn (imat+2, imat+1) + 0.5d0 * f3yizi
dyn (imat+2, imat+2) = dyn (imat+2, imat+2) + 0.5d0 * f3zizi

480    continue
490    continue
500    continue
    tote = tote2 / 2.d0 + tote3 / 6.d0

    print *, 'Total Energy: E/N=', tote / np

    return
    end

C******************************************************************************