Modeling Multi-Band Effects of Hot-Electron Transport in Simulation of Small Silicon Devices by a Deterministic Solution of the Boltzmann Transport Equation Using Spherical Harmonic Expansion

Surinder Pal Singh

Doctoral Dissertation
1998
Department of Electrical Engineering
University of Maryland
College Park, MD 20742

MODELING MULTI-BAND EFFECTS OF HOT-ELECTRON TRANSPORT IN SIMULATION OF SMALL SILICON DEVICES BY A DETERMINISTIC SOLUTION OF THE BOLTZMANN TRANSPORT EQUATION USING SPHERICAL HARMONIC EXPANSION

by

Surinder Pal Singh

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1998

Advisory Committee:

Professor Isaak D. Mayergoyz, Chairperson/Advisor
Professor Neil Goldsman, Co-Advisor
Professor Martin C. Peckerar
Professor Theodore J. Rosenberg
Professor Leonard S. Taylor

ABSTRACT

Title of Dissertation: MODELING MULTI-BAND EFFECTS OF
HOT-ELECTRON TRANSPORT IN
SIMULATION OF SMALL SILICON
DEVICES BY A DETERMINISTIC
SOLUTION OF THE BOLTZMANN
TRANSPORT EQUATION USING
SPHERICAL HARMONIC EXPANSION

Surinder Pal Singh, Doctor of Philosophy, 1998

Dissertation directed by: Professor Isaak D. Mayergoyz
Department of Electrical Engineering

Solution of Boltzmann equation by a spherical-harmonic expansion approach is a computationally-efficient alternative to Monte Carlo. In this dissertation we extend this technique to compute the distribution function in multiple bands of silicon, using a multi-band band-structure which is accurate for high energies. A new variable transformation is applied on the spherical harmonic equations. This transformation (a) improves the numerical properties of the equations by enhancing the diagonal dominance of the resulting equations; (b) accounts for exponential dependence of the distribution function on energy as well as electric potential; and (c) opens the possibility of using superior Poisson solvers (d) while retaining the linearity of the original equations intact. The resulting Boltzmann equations are discretized using the current-conserving control-volume approach. The discretized equation are solved using line successive-over-relaxation (SOR) method. Numerical noise in the distribution was analyzed to be originating from the absence of coupling. Noise is removed by using acoustic phonons in inelastic approximation. A novel self-adjoint easy-to-discretize formulation for the inelastic acoustic phonons is developed. A test case of thermal equilibrium for multi-band is derived and used to validate the code. Hole-continuity and Poisson equation were solved along with the multi-band Boltzmann equations. The equations are solved in a Gummel-type decoupled loop. A n^+nn^+ device is simulated to test the simulator. The simulator is then applied to study a one-dimensional short-base bipolar junction transistor. While these simulations are self-consistent, a two-dimensional sub-micron MOSFET is simulated in a non-self-consistent manner.

Key Words: Device Simulation, Spherical Harmonic Expansion, Boltzmann Transport Equation, Semi-Classical Transport, Monte Carlo, Computer-Aided Design (CAD), Technology Computer-Aided Design (TCAD), Non-Equilibrium Transport, Non-Local Transport, Hot Electron, Multi-Band Band-Structure, Device Physics, Computational Methods in Semiconductors, Numerical Methods, Control-Volume Discretization, Sub-Micron MOSFET, Short-Base Bipolar Junction Transistor (BJT), n^+nn^+ Device.

© Copyright by
Surinder Pal Singh
1998

NOTE FOR CITATION

This copy of the dissertation has been made for personal distribution. It differs from the official version in typography—the most notable difference being that this one is single-spaced while the Graduate-School version is double-spaced. Content-wise, of course, this dissertation is identical to the official version. A reader wishing to cite material in this dissertation has two possibilities: avoid references to page numbers, but instead cite chapter numbers, section numbers, equation numbers, figure numbers, or table numbers; alternatively, the reader may obtain the official version. The official version can be obtained from any of the following sources: author, dissertation advisor, the University of Maryland library, or UMI dissertation service.

DEDICATION

To my Parents:
Sudarshan Kaur and Prabjit Singh
And
My Teachers

TABLE OF CONTENTS

LIST OF TABLES xvi

LIST OF FIGURES xviii

PREFACE xxiii

ACKNOWLEDGEMENT xxvii

1 INTRODUCTION 1

1.1 Chapter Introduction 1
1.2 Motivation for Device Simulation 1
1.3 Device Simulation Approaches 2
  1.3.1 Drift-Diffusion 4
  1.3.2 Hydrodynamic or Energy-Transport 4
  1.3.3 Monte Carlo 4
1.4 Device Simulation by Spherical Harmonics 5
1.5 Literature Review of Spherical Harmonic Approach 6
1.6 Scope & Contributions of the Dissertation 8
1.7 Organization of the Dissertation 9
1.8 Chapter Summary 9

2 SPHERICAL HARMONIC EXPANSION OF BOLTZMANN TRANSPORT EQUATION: DERIVATIONS AND FUNDAMENTALS 11

2.1 Chapter Introduction 11
2.2 Boltzmann Transport Equation 12
2.3 Spherical Harmonic Expansion of the Boltzmann Equation . . 14
2.4 Average Quantities 19
  2.4.1 Basics 19
  2.4.2 Summation Over k-Space 19
  2.4.3 Integration Over k-Space 20
  2.4.4 Electron Concentration 22
  2.4.5 Electron Current and Velocity 23

2.5 Collision Terms . . . 26
    2.5.1 Optical Phonon Scattering . . . 27
    2.5.2 Inter-Band Optical Phonon Scattering . . . 28
    2.5.3 Inelastic Acoustic Phonon Scattering . . . 29
    2.5.4 Ionized Impurity Scattering . . . 30
    2.5.5 Impact Ionization Scattering . . . 31
2.6 Final Form of the Boltzmann Equation . . . 34
2.7 Boundary Conditions . . . 35
    2.7.1 Ohmic Boundary Condition . . . 35
    2.7.2 Insulator Boundary Condition . . . 36
    2.7.3 Energy Boundary Condition . . . 37
2.8 Chapter Summary . . . 40

3 MULTI-BAND BAND-STRUCTURE 43
3.1 Chapter Introduction . . . 43
3.2 Multi-Band Band-Structure . . . 44
3.3 Boltzmann Equation in Multi-Band Band-Structure . . . 46
    3.3.1 Total Distribution . . . 48
3.4 Collision Term and Transport Model . . . 48
    3.4.1 Optical Phonon Scattering . . . 48
    3.4.2 Inter-Band Optical Phonon Scattering . . . 49
    3.4.3 Acoustic Phonon Scattering . . . 49
    3.4.4 Ionized Impurity Scattering . . . 49
    3.4.5 Impact Ionization Scattering . . . 50
3.5 Chapter Summary . . . 50

4 TRANSFORMATION AND DISCRETIZATION OF THE BOLTZMANN TRANSPORT EQUATION 53
4.1 Chapter Introduction . . . 53
4.2 Transformation . . . 54
    4.2.1 The Pre-Maxwellian Variable . . . 54
    4.2.2 Boltzmann Equation in Pre-Maxwellian Variable . . . 55
    4.2.3 Collision Terms in Pre-Maxwellian Variable . . . 56
        4.2.3.1 Optical Phonon Scattering . . . 56
        4.2.3.2 Inter-band Optical Phonon Scattering . . . 57
        4.2.3.3 Inelastic Acoustic Phonon Scattering . . . 57
        4.2.3.4 Ionized Impurity Scattering . . . 59
        4.2.3.5 Impact Ionization Scattering . . . 59
    4.2.4 Final Equation in Pre-Maxwellian Variable . . . 60
4.3 Discretization . . . 61
    4.3.1 Principles of Control-Volume Discretization . . . 62

4.3.2 Discretization of the Boltzmann Equation 63
4.3.3 Discretization of Collision Terms 67
4.3.3.1 Optical Phonon Scattering 67
4.3.3.2 Inter-Band Optical Phonon Scattering 67
4.3.3.3 Inelastic Acoustic Phonon Scattering 68
4.3.3.4 Impact Ionization Scattering 68
4.3.4 Putting it Together 68
4.3.5 Boundary Conditions 69
4.3.5.1 Ohmic and Insulator Boundary Condition . . 70
4.3.5.2 Energy Boundary Condition 70
4.4 Advantages of Pre-Maxwellian Variable 71
4.4.1 Variation of Distribution in Energy and Space 73
4.4.2 Diagonal Dominance in Pre-Maxwellian Variable 73
4.4.2.1 Original Variable 74
4.4.2.2 Pre-Maxwellian Variable 76
4.4.2.3 Discussion 77
4.4.2.4 Alternate Notation 78
4.4.2.5 Why Pre-Maxwellian Variables Give Diagonal Dominance? 78
4.4.3 Poisson-Equation Solvers with Pre-Maxwellian Variable 79
4.5 Chapter Summary 80

5 SIMULATION OF N$^+$NN$^+$ DEVICE STRUCTURE 81
5.1 Chapter Introduction 81
5.2 Solution of Discretized Equations 82
5.3 Solution of the Coupled Equations 83
5.4 The Simulator 85
5.5 Simulation of a n$^+$nn$^+$ Device 85
5.5.1 Device Structure & Grid 87
5.5.2 Simulation Results at Various Biases 87
5.5.3 Drift-Diffusion versus Boltzmann Equation 88
5.5.4 Distribution Function at 5V 89
5.6 Thermal Equilibrium 89
5.7 Convergence of Numerical Solution 90
5.8 Numerical Noise and Inelastic Acoustic Phonons 91
5.8.1 Summary of the Idea 91
5.8.2 Detailed Explanation 92
5.9 Chapter Summary 93

6 SIMULATION OF BIPOLAR JUNCTION TRANSISTOR 107
6.1 Chapter Introduction . . . 107
6.2 Introduction to Bipolar Junction Transistor . . . 108
6.3 Simulation of the Bipolar Junction Transistor . . . 109
  6.3.1 Base Boundary Condition . . . 109
6.4 Simulation Results . . . 109
  6.4.1 Device Structure & Doping . . . 109
  6.4.2 Grid & Simulation Time . . . 110
  6.4.3 Electric Potential & Field . . . 110
  6.4.4 Distribution Function . . . 110
  6.4.5 Electron Velocity & Concentration . . . 111
  6.4.6 Current Density . . . 112
  6.4.7 Hole Concentration . . . 112
  6.4.8 Average Energy . . . 112
  6.4.9 Impact Ionization . . . 112
  6.4.10 Effect of Including Impact Ionization . . . 113
  6.4.11 Effect of Including Ionized Impurity . . . 113
  6.4.12 Comparison with Drift-Diffusion . . . 114
  6.4.13 Ballistic Transport . . . 114
  6.4.14 Graded Doping in Base . . . 115
  6.4.15 Thermal Equilibrium . . . 115
6.5 Chapter Summary . . . 116

7 SIMULATION OF MOSFET 133
7.1 Chapter Introduction . . . 133
7.2 Introduction to MOSFET's . . . 134
7.3 Numerics for MOSFET Simulation . . . 135
  7.3.1 Boundary Conditions . . . 135
  7.3.2 Two Space Dimensions . . . 135
  7.3.3 Decoupled Solution Technique . . . 136
  7.3.4 Grid Generation . . . 136
7.4 The Simulator . . . 137
7.5 Results . . . 138
  7.5.1 Device Structure & Doping . . . 138
  7.5.2 Grid & Simulation Time . . . 138
  7.5.3 *I*-*V* Curves . . . 139
  7.5.4 Distribution Function . . . 140
  7.5.5 Average Quantities . . . 141
  7.5.6 Ballistic Transport . . . 142
  7.5.7 Thermal Equilibrium . . . 143
  7.5.8 Effect of Channel-Length . . . 143

7.6 Chapter Summary 144

8 CONCLUSION 171

A CURRENT-CONSERVATION PROPERTY OF THE SPHERICAL-HARMONIC BOLTZMANN EQUATION 173

A.1 Proof of Current-Continuity 173
A.2 Discussion 175

B MULTI-BAND BOLTZMANN EQUATION AT THERMAL EQUILIBRIUM 177

C HARMONIC-MEAN SCHEME 179

C.1 Derivation 179
C.2 Similarity to Scharfetter-Gummel Discretization 181
C.3 Similarity to Liang-Goldsman-Mayergoyz Scheme 182

D POISSON EQUATION 183

D.1 The Boundary-Value Problem 183
  D.1.1 Boundary Conditions 184
D.2 Discretization 184
  D.2.1 Damping Term 186
  D.2.2 Final Form 186

E HOLE-CONTINUITY EQUATION 189

E.1 The Boundary-Value Problem 189
  E.1.1 Boundary Conditions 190
E.2 Scharfetter-Gummel Discretization 190

F LIST OF SYMBOLS 193

BIBLIOGRAPHY 197

INDEX 209

LIST OF TABLES

3.1 Numerical values for the band-structure 46
3.2 Values of transport parameters 49

4.1 Diagonal dominance of the discretized Boltzmann equation . . . 78

7.1 Numerical values of the various parameters for the simulated MOSFET 145

LIST OF FIGURES

1.1 Hierarchy of computer-aided design (CAD) in VLSI design. . . . 3
1.2 Relative computational and physical rigor of various transport models 7

2.1 Ludwig Boltzmann. Photograph and life. 13
2.2 Coordinate systems used in the spherical-harmonic expansion . 16
2.3 Illustration of the impact ionization process 33
2.4 Derivation of the energy boundary condition 40
2.5 Alternate view of energy boundary condition 41

3.1 The multi-band band-structure 45
3.2 Density of states for multi-band band-structure 47
3.3 Optical and acoustic phonon scattering rate. 50
3.4 Ionized impurity scattering rate 51
3.5 Impact ionization scattering rate 52

4.1 Schematic of the grid on the domain of the spherical-harmonic Boltzmann equation 64
4.2 Grid for the solution of the spherical-harmonic Boltzmann equation 65
4.3 The control-volume at an insulator or ohmic boundary 70
4.4 Application of the energy boundary condition on the grid 72

5.1 Illustration of the line-SOR approach to solve the Boltzmann equation 84
5.2 Gummel's self-consistent decoupled scheme 86
5.3 Structure of the simulated n^+nn^+ device 94
5.4 Electric potential in n^+nn^+ device for three applied biases . . . 95
5.5 Electric-field profile within the n^+nn^+ device 95
5.6 Average electron velocity in the n^+nn^+ device 96
5.7 Electron concentration in the n^+nn^+ device 96
5.8 Current density in the n^+nn^+ device 97
5.9 Average electron energy in the n^+nn^+ device 98
5.10 Impact ionization coefficient in the n^+nn^+ device 98

5.11 Electron concentration in a n^+nn^+ device by drift-diffusion and Boltzmann equation . . . 99
5.12 Electric field in the n^+nn^+ device by drift-diffusion and Boltzmann equation . . . 99
5.13 Distribution function in the n^+nn^+ device at 5V . . . 100
5.14 The distribution function in n^+nn^+ device at thermal equilibrium 101
5.15 Electric field in the n^+nn^+ device at thermal equilibrium . . . 102
5.16 Convergence behavior of the self-consistent Gummel scheme . . 103
5.17 Convergence behavior of various matrix solutions of Boltzmann equation . . . 104
5.18 Noisy distribution function obtained by considering acoustic phonons in elastic approximation . . . 105
5.19 Non-noisy distribution function obtained by considering acoustic phonons in elastic approximation . . . 106

6.1 Two-dimensional structure of a planar n-p-n integrated-circuit bipolar junction transistor . . . 117
6.2 doping profiles for the n-type and p-type BJT's . . . 117
6.3 Electric potential and electric field for the n-type and p-type BJT's . . . 118
6.4 Electron energy distribution function in the n-type BJT . . . 119
6.5 Electron energy distribution function in the p-type BJT . . . 120
6.6 Electron velocity and concentration profile in the n-type and p-type BJT . . . 121
6.7 Electron current density in the n-type and p-type BJT's . . . 122
6.8 Hole concentration profile in the n-type and p-type BJT's . . . 122
6.9 Average electron energy profile in the n-type and p-type BJT's . 123
6.10 Profile of impact ionization coefficients in n-type and p-type BJT's 124
6.11 Effect of including impact ionization on the distribution function 125
6.12 Effect of including ionized impurity scattering on the electron velocity . . . 126
6.13 Comparison of electron concentration calculated by drift-diffusion and Boltzmann equation . . . 126
6.14 Ballistic transport in the base-collector region of n-type BJT: line plot at three locations . . . 127
6.15 Ballistic transport in the base-collector region of n-type BJT: contour plot . . . 128
6.16 Effect of graded base doping profile on electron velocity . . . 129
6.17 Distribution function in the n-type BJT at thermal equilibrium 130
6.18 Average electron energy and electric field in the n-type BJT at thermal equilibrium . . . 131

7.1 Schematic of a conventional MOSFET 146
7.2 Flow chart for MOSFET simulation 147
7.3 Grid in the simulated MOSFET 148
7.4 Doping profile of the simulated MOSFET 149
7.5 The simulated transfer characteristics of the MOSFET 150
7.6 The simulated output characteristics of the MOSFET 151
7.7 Simulated substrate current in the MOSFET 151
7.8 Current crossing through a vertical cross-section in the MOSFET 152
7.9 Electric potential in the MOSFET 152
7.10 Vector plot of electric field within the MOSFET 153
7.11 Distribution function along a horizontal section of the MOSFET under the oxide 154
7.12 Distribution function along a horizontal section of the MOSFET at source/drain junction depths 155
7.13 Distribution function along a horizontal section in the MOSFET deep into the substrate 156
7.14 Electron concentration in the MOSFET as computed by drift-diffusion and Boltzmann equation 157
7.15 Vector plot of the current density in the MOSFET 158
7.16 Electron velocity in the MOSFET 159
7.17 Average energy of electrons in the MOSFET 160
7.18 Generation rate due to impact ionization in the MOSFET . . . 161
7.19 One-dimensional plots of doping profile, electric potential, lateral electric field, average energy under the oxide of the MOSFET 162
7.20 One-dimensional plots of average electron velocity, electron concentration, impact-ionization generation rate under the oxide of the MOSFET 163
7.21 Electron concentration in the various bands under the oxide of the MOSFET 164
7.22 Ballistic Transport in the MOSFET 165
7.23 Distribution function in the MOSFET at thermal equilibrium . 166
7.24 Electric field and average energy at thermal equilibrium in MOSFET 167
7.25 Effect of channel-length on MOSFET operation 168
7.26 Simulated transconductance of the MOSFET's as a function of effective channel-length 169

C.1 Grid for derivation of the harmonic-mean scheme 180

D.1 Grid for the discretization of the Poisson equation 185

E.1 Grid for the discretization of the hole-continuity equation 190

PREFACE

During my third year at IIT (Indian Institute of Technology), Kanpur, India, in the Fall of 1987, I took the analog circuit design course under Prof. R. N. Biswas. Aesthetic beauty of the analog circuits attracted me to the field of semiconductors; so when I was looking for an area to concentrate for my Masters at the University of Maryland, I chose Microelectronics. In 1992 Spring, my final semester of my MSEE degree, I took EE648 O, "Numerical analysis of Semiconductors," from Prof. Isaak Mayergoyz. This course was my first introduction to device simulation. Two semesters later I joined him as a graduate student.

Device simulation is a very interesting and important field. It requires skills spanning over many areas of science and engineering: electrical engineering, physics, computer science, mathematics, numerical analysis, material science amongst many. The field is young enough that the early pioneers are still around. In fact, I had the good fortune to meet at Intel one of the "Gods" of device simulation: Don Scharfetter.

The famous quotation by R. W. Hamming, [1] "The purpose of computing is insight, not numbers," is valid for device simulation. The utility of device simulation is the insight it provides into device operation, which is then used to improve device performance.

Developing device simulators involves developing the numerics as well as writing the code. Initially, writing and debugging the simulator code takes up all the effort; the pleasure of seeing the simulator reveal insight into devices comes rather late. I used to sum this in the following statement, which is a slightly-modified version of a lament I read from an experimentalists: [2] "Numerical device simulation is what we do in between debugging our code." The development of the simulator is made difficult because sometimes one is not sure whether the simulator is not working because of weak numerics, or due to presence of program bugs. This is exacerbated by the fact that journals publish only successful simulations; we usually do not encounter papers that present a failed numerical scheme. During the course of research I did both things: I have discarded many promising numerical schemes because they did not seem to work when coded; and I have sometimes pursued some numerical ideas of dubious merit too vigorously.

---

[1] R. W. Hamming, *Numerical Methods for Scientists and Engineers*, McGraw Hill Book Company, New York, 1973.

[2] "Experiments is what we do in between repairing instruments." Anonymous.

Despite the many baby's I threw with the bath water and many rat holes I ended in, I have immensely enjoyed my dalliance with this subject. One such incident is the idea for the pre-Maxwellian variables discussed in this thesis. In the summer of 1995, I was perturbed by my self-consistent simulator not converging to the right answer. Every trick in my bag failed. I was cogitating on the cause when, while sitting in the back of the hall and listening to the *kirtan* music of Bhai Jagmohan Singh, the idea for the pre-Maxwellian dawned on me. It was a reincarnation of an idea I had once explored with Prof. Mayergoyz. I went to work on this idea; it looked good on paper, but the code did not work. I shelved the idea at that time; little did I know that I would be revisiting this idea a year and half later, and it will form the basis of all simulations of the thesis.

This dissertation brings to close my academic life as a student. I have attempted to make this document as readable and usable to my readers. This is at two levels: (a) content chosen to be written in the thesis; and (b) the writing style of the presentation. On the content front, I have shied away from regurgitating derivations that are already present in literature, referring to appropriate citations instead. In addition, I relegated many derivations to the appendices. I have striven to achieve the dual aim of providing enough detail so that a researcher, like a graduate student, has enough details to reproduce my results, as well as keep the advanced reader interested. On the writing-style front, I have tried to make the presentation crisp, direct, clear, pointed, unambiguous, and vigorous. I have tried to eschew from creating acronyms, and kept their use to the minimum, using only the most common of all (like MOSFET, BJT). I have added an index for easy reference. I have been tremendously aided by books on good writing style in general [3] and writing for science and engineering. [4] Two publications, for me, have been the paragon of good writing-style: Prof. Mayergoyz' Poisson equation paper [95], and Prof. Patankar's book [113]. Despite my efforts, there could be many spelling, grammar, and stylistic errors remaining in the dissertation; I apologize to the readers for that.

---

[3] D. Hacker, *A Writer's Reference*, Bedford Books of St. Martin's Press, Boston, MA, 1995.

[4] J. Kirkman, *Good Style Writing for Science and Technology*, E & FN Spon, London, UK, 1992.

Santa Clara, California Surinder Pal Singh
April 1998

ACKNOWLEDGEMENT

While it is quite common for all graduate students to be thankful to their advisors, my debt of gratitude goes beyond. Without the help of my advisor, Prof. Isaak Mayergoyz, there was no hope for me to ever finish this dissertation. In Maryland, and in California, he has provided unprecedented support and guidance. He has trained me in device simulation, first in his graduate class, and then as his graduate student. In my years of association with him I have come to realize that his sharp intellect is but only *one* component of him. He put up with me—a difficult, if not impossible, student—and provided very deep encouragement. Not only have I learned research from him, I have also become a better person by knowing him. The completion of this dissertation owes more to him, than I can ever realize. My only regret is that I could have learned more from him.

Prof. Neil Goldsman, as the co-advisor, provided intensely creative guidance. It has been a pleasure and rare privilege to learn device simulation and device physics from the pioneer of spherical-harmonic device-simulation approach. He was always willing to give his time and energy to my problems, however minor. I owe him a great debt of gratitude. Prof. Goldsman has shown great patience in dealing with a difficult student like me; for this, I shall always be humbly and sincerely grateful. Both Prof. Goldsman and Prof. Mayergoyz have overlooked many of my deficiencies, perhaps more than I realize.

I acknowledge SRC (Semiconductor Research Corporation) and NSF (National Science Foundation) for financial support through the research grants.

I also wish to thank the members of the defense committee: Prof. T. J. Rosenberg, Prof. L. S. Taylor, Prof. M. C. Peckerar. It is indeed a happy happenstance that Prof. Rosenberg should serve as Dean's representative—he was my Master's thesis advisor, and I decided to join the Ph.D. program because of his encouragement.

Dr. Kartik Ramaswamy was a great pillar of help to encourage me in this dissertation. He would regularly call me from Austin to gently egg me to finish. He continued this encouragement when he moved to Santa Clara. He has often patiently checked me from getting carried over. Being quite familiar with my research he was able to offer many ideas, especially when the chips were down. I acknowledge the tremendous help given by Dr. Rajiv Madabhushi. He has offered solution to problems encountered while coding the simulators and writing of the dissertation. Not only did he give me his drift-diffusion code for MOSFET's, he would consistently inquire about the progress of the dissertation. I also owe to him many evenings of discussions on Jack Nicholson and Marlon Brando. Kartik and Rajiv were indispensable for me.

In the laboratory at Maryland I have the great fortune to befriend Tse "Henry" Chun. I have enjoyed my deep conversations with him. I cannot express enough gratitude for his help in administrative matters in Maryland, since I have been in California. To Dr. Rajat Garg, my roommate in Maryland, I owe the pleasure of many hours of fruitful discussion on CFD (computational fluid dynamics). Many sections in this dissertation bear the mark of those discussions. At his graduation from Stanford, I had privilege of meeting Prof. Joel Ferziger, a recognized expert in CFD. I am thankful to him for looking at my problem and offering insight. I am extremely grateful to Dr. P. M. de Zeeuw for pointing out the similarity of Scharfetter-Gummel discretization with the harmonic-mean scheme [106].

In my numerous courses I learned from the one of the finest teachers. At the University of Maryland I had Profs. Linda Milor, Martin Peckerar, Kazuo Nakajima, Neil Goldsman, Isaak Mayergoyz as my teachers. At IIT (Indian Institute of Technology), Kanpur, I had the fortune of studying from Profs. R. N. Biswas, Raminder Singh, Vijay Gupta, R. Raghuram, K. R. Sarma, Pradip Sircar, Vijay Deshpande, and P. N. Rastogi amongst others.

Graduate school at Maryland was vibrant; I had the opportunity to meet many talented graduate students: Drs. Ken Hennacy, Gwo-Chung Tai, Dan Kerr, John Stanley, Wenchao Liang, Yu-Jen Wu, Amr Adly, Lindor Henrickson, Shane Dias. I also benefited immensely from many other students: Badri Gopalan, Satpal Singh, Dr. Scott Meritt, Dr. John Paquette, Dr. Pat Mead.

Life in Maryland was richer by my interaction with many friends. My roommates provided a very nice environment to stay, for this I thank Drs. Sanjay Khare, Prosenjit Mazumdar, Milind Rane, Muthukumara Swamy Mani, Harmeet Singh, Rajat Garg. I have enjoyed numerous dinners and Planet X coffees with Pritiraj Mohanty and Sanjay Khare. Friends like Subir Bannerji, Sanjeev Munjal, Dr. Mahendra Verma, Alok Bharadwaj, Deven Verma have shown me how to be a better person. I would like to thank Gurdeep Singh, and Ravinder Singh for guidance. To Dr. Jaspreet Arora I owe many dinners and movies. I thank Bhai Jagmohan Singh for the weekly kirtan music.

I would like to appreciate the editors and readers of the campus newspaper, *The Diamondback*, for the enjoyable time I had drawing editorial cartoons; the McKeldin library for the newspapers I read; and the Hornbake library for all the Western movies I checked out.

I would like to thank everyone for the help I received from the world-village community of newsgroup readers: comp.graphics.apps.gnuplot, especially Dick Crawford (all figures in this dissertation were made in gnuplot); comp.text.tex (this dissertation was typeset in latex); the Linux newsgroups (all computer work was done a PC with Linux operating system); and sci.physics.computational.fluid-dynamics for many numerical tips. I am grateful to Prof. Edmund R. Robertson and Prof. John O'Connor of the University of St. Andrews, Scotland, the maintainers of the web-site [11], for permission to use the photograph of Ludwig Boltzmann (page 13).

The challenging work environment at Intel Corporation has provided great learning for me. Specifically, I would like to thank the members of my first project at Intel: Sanjay Natu, Melik Isbara, Micah Barany, Raju Abhyankar, Sarah Bates and Arvind Purushotham. I acknowledge many enjoyable and informative discussions on design, technology, and analog circuits with Gerald Pasdast, Dr. Brian Cherkauer, John Orton, Chih-Hung Chung, and Bal Sandhu.

I finally thank my brother Maj. Manmohan Singh Vilkhu, and my parents, to whom the dissertation is dedicated, for the motivation towards education and fostering talent.

In the end, I would like to specifically thank Prof. P. N. Rastogi. Not only did he teach me "Social Cybernetics," he also gave me a glimpse into the world of knowledge: Vast and beautiful, enchanting and peaceful.

CHAPTER 1

INTRODUCTION

1.1 Chapter Introduction

This chapter is an introduction to the entire dissertation. This chapter is organized as follows: Section 1.2 outlines the motivation and need for technology computer-aided design (TCAD) tools in the semiconductor industry. Various approaches to device simulation—like drift-diffusion, hydrodynamic, and Monte Carlo—are discussed in Section 1.3 on the next page, followed by a discussion on spherical-harmonic approach in Section 1.4 on page 5. The spherical-harmonic literature is briefly reviewed in Section 1.5 on page 6. The scope and contributions of this dissertation are discussed in Section 1.6 on page 8. Finally, the organization of the dissertation is outlined in Section 1.7 on page 9.

1.2 Motivation for Device Simulation

Semiconductors, sometimes referred to as the *rice* of the industry [1], have made tremendous advances in the recent decades. To increase functionality, the number of transistors on a chip have increased at an exponential rate; in fact, according to the Moore's Law, number of transistors on a chip doubles every eighteen months. Logic chips, like microprocessors, currently can easily have 5–15 million transistors; in the year 2003 this number is expected to be 76 million; in the year 2006 it is expected to be at 200 million; while it might be as large as 520 million in year 2009 [2].

Minimum feature size on an integrated circuit is continuously shrinking. This is because smaller feature sizes have many advantages: higher transistor counts, which means more functionality; lower power; faster speed of operation. While decades ago the minimum channel-length of metal-oxide-semiconductor field-effect transistor (MOSFET) was 20μm, but the current state-of-the-art in production is 0.25–0.15μm. The industry organizations are already in the process of developing 0.1-μm technology, which is expected to ship in the year 2003. It is estimated that MOSFET channel-length of products shipped in the year 2009 will reach 0.05μm [2]. Technology is expected to continue shrinking in the near future.

Since the modern chip design and manufacturing is an extremely complex and involved process, computer-aided design (CAD) tools are used in all stages [72, 81, 82]. In the semiconductor industry there are two categories of CAD tools: electronic circuit CAD (ECAD); and technology CAD (TCAD) [3–8].

ECAD, which is sometimes also referred to as "higher-level CAD" [5], provides software tools to the chip designers in the areas of logic and circuit design, layout, design-rule checking, timing analysis, floor-planning, hardware description languages, silicon compilers. It is primarily associated with product design in relatively mature production technology.

TCAD, on the other hand, typically is applied during the development of the technology. It ranges from equipments simulation, process simulation, device simulation, and sometimes even small-circuit simulation. Equipment simulation predicts the behavior of equipment used in a fabrication plant. Process simulation would then simulate of each step of the chip manufacturing to predict the final device structure and doping profiles. This information is in turn used in device simulation to predict the external and internal electrical behavior of the devices like I-V curves, electric fields, electron and hole energies and concentrations [3]. The predictive nature of TCAD allows technologists to develop new processes and devices in a faster and resource-efficient manner [4].

While the relatively-more-mature field of ECAD fuels its development mainly from the algorithms and techniques of computer science, TCAD, on the other hand, culls from the academic disciplines of applied mathematics, physics, chemistry, chemical engineering, electrical engineering, material science, and computer science [5].

Fig. 1.1 on the facing page shows the hierarchy of CAD tools used in the very large-scale integrated (VLSI) circuit design [5]. Information feeds upward, from the primitive to the higher level: equipment simulation → process simulation → device simulation → circuit simulation and higher. The focus of this dissertation is on this technologically-important aspect of TCAD: *Device Simulation.*

1.3 Device Simulation Approaches

The operation of semiconductor device can be modeled by a set of transport equations along with the Poisson equation, The transport equations, which are typically coupled partial differential equations, are not unique; they depend

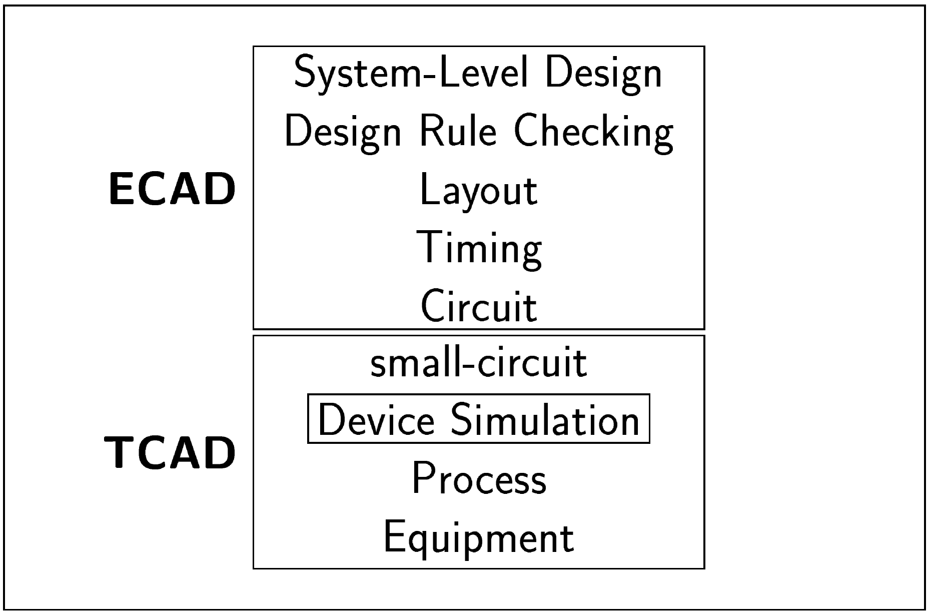


Figure 1.1: Hierarchy of computer-aided design (CAD) in semiconductor design and manufacturing. Electronic CAD (ECAD) is used for chip design; technology CAD (TCAD) is used for technology development. *Device simulation*—which is the subject of this dissertation—forms a very important part of this process. After [5].

on the level of accuracy desired. These equations can be solved to understand and predict devices in two ways: simplified closed-form solution or rigorous numerical solution.

Closed-form analytical solutions are obtained by approximating the governing transport equations on simplified piece-wise linear doping profiles. Rigorous numerical methods, on the other hand, solve these partial differential equations exactly and rigorously by numerical techniques: The device can be divided into small regions by grids, and the equations solved by finite-difference or finite-element methods. This allows flexibility in having arbitrary device shapes and doping profiles, and ease of incorporating more physically sophisticated models. Unfettered by blanket assumptions, the numerical simulation of devices gives reliable and accurate solutions. In the absence of closed form solutions, however, the physical meaning of device parameters and their influence on device performance may sometimes be obscured [54]. *Numerical device simulation, therefore, can be defined as a numerical solution of semiconductor transport equations which include relevant physical effects in a realistic device structure.*

Very intimately connected with the simulation is the choice of the governing transport model equations. A sophisticated model may include more physics and may predict device behavior more accurately; although it may consume more computational resources. A simpler model, on the other hand, may be computationally efficient; although the model may describe the device behavior less accurately. Following models, in order of increasing complexity, are used in industry for device simulations:

1. Drift-Diffusion
2. Hydrodynamic or Energy-Transport
3. Monte Carlo

Following sub-sections discuss these models.

1.3.1 Drift-Diffusion

Drift-diffusion is the oldest models of semiconductor transport [72]. The model gets its name by the assumption that the charge transport of the carriers (electrons and holes) is a result of *drift* (due to electric fields) and *diffusion* (due to the concentration gradients). This is the most common transport model. Sophisticated techniques exist for solving this model. Its major strengths are (a) mathematical simplicity and robustness, (b) modest CPU (central processing unit) demands, and (b) easy calibration. On the down side, however, the model is not accurate for devices where the electric is large, or rapidly varying—typical of modern sub-micron MOSFET and bipolar transistors [9, 10]. This is due to the fact that the drift-diffusion model does not model hot-electron transport.

1.3.2 Hydrodynamic or Energy-Transport

The hydrodynamic model assumes that, in addition to the drift and diffusion, the current density has a thermal component to it [74]. It assumes that carriers can be assigned a *temperature*, which is different from the lattice temperature [56, 62, 74]. The advantage of the hydrodynamic device simulation is that it has more physical detail. On the flip side it is (a) numerically unstable, (b) requires more computational time, and (c) is not as easy to calibrate.

1.3.3 Monte Carlo

Monte Carlo technique solves the Boltzmann transport equation [62–70, 87–90, 92]. It is a *stochastic* technique which generates random numbers to microscopically simulate electron motion—drift in electric field and collision with lattice or other electrons. Monte Carlo approach can successfully solve the multi-dimensional Boltzmann equation, something that is not a trivial task for direct numerical techniques. It can include, with relative ease, complex band-structure and scattering processes. The Monte Carlo technique, however, is encumbered by intensive computational demands—large CPU (Central Processing Unit) times and large memory requirements. This becomes more severe when the Monte Carlo simulation has to be self-consistent with the Poisson Equation. Monte Carlo method consumes even more computation time for devices which have regions of high doping, low electric fields, and retarding barriers [67, 89]—which, of course is true for most devices, especially MOSFET's and bipolar junction transistors. This is inherent in the nature of the Monte Carlo technique. A very large number of particles are needed to obtain a *statistically significant* estimate, free from *statistical noise*. This statistical noise manifests in two ways. Firstly, the *tail* (high-energy) region of the distribution function is noisy and inaccurate, especially at low electric field—devices almost always have large regions of low electric fields. Secondly, the electric currents predicted by Monte Carlo at low applied bias are very noisy.

1.4 Device Simulation by Spherical Harmonics

We make conflicting demands on device simulation TCAD: On one hand we want more physical rigor to model the modern sub-micron devices; on the other, we would desire computational efficiency and numerical robustness. Spherical-harmonic approach is one such solution technique which promises to deliver both. Fig. 1.2 on page 7 compares the spherical-harmonic approach with the others. Physics in the spherical-harmonic approach is much more than that in drift-diffusion and hydrodynamic models; but it is comparable to that in Monte Carlo. The computation time is, on the other hand, is much less that for Monte Carlo. Chang, *et al.*, [44] report the computation time of 10 seconds for spherical-harmonic simulation of a bipolar junction transistor (BJT), while it took 10 hours for Monte Carlo. Spherical-harmonic approach, therefore, is a computationally efficient alternative to the Monte Carlo for the solution of the Boltzmann equation [17–51]. It is a direct *deterministic* solution, as compared the stochastic solution of the Monte Carlo approach. It does not, therefore, suffer from the statistical noise in the tail of the distribution function. Unlike drift-diffusion it incorporates non-local non-equilibrium effects. Unlike drift-diffusion and hydrodynamic, this approach provides the distribution function for the carriers, which in turn allows accurate computation of distribution-function-dependent hot-electron effects like gate and substrate currents in MOSFET's [84, 85]. Hot electrons, incidently, are those electrons in the semiconductor that have much higher kinetic energy than the average carrier population [85].

The basic idea of spherical-harmonic approach is to expand the distribution function in momentum space using spherical harmonics. In fact, this expansion reflects the spherical symmetry inherent due to the spherical symmetry of the band-structure and the randomizing nature of most scattering processes [17]. In the extreme case, the equilibrium distribution function is fully spherically symmetric. The spherical-harmonic basis function exploits this symmetry. The expansion reduces the dimensionality in momentum space and allows for analytical evaluation of the scattering integral.

Spherical-harmonic-expansion approach can also be interpreted as a *spectral* method, similar to those used in the discipline of computational fluid dynamics (CFD) [114, 115]. By expanding the distribution in spherical harmonics, we have used a spectral method in azimuthal- and polar-angle direction, and the usual finite-difference in radius (energy) and space. Spectral methods are well-known for their high accuracies arising from low discretization error [115].

1.5 Literature Review of Spherical Harmonic Approach

Spherical harmonics, or their simplified cousins Legendre Polynomials, were first introduced in solid-state physics in 1950's [49]. In recent years, spherical-harmonic approach to solving the Boltzmann equation was proposed by Prof. Neil Goldsman in 1989 in his seminal work in [20, 47]. That aroused considerable interest in the device-simulation community and attracted many research groups to this area; since then a considerable body of literature has built up: [17–28, 30–43, 46–51] At the time of writing this dissertation, there are three major groups active in this area: (1) Italian group at University of Bologna [21, 27–31, 35–39, 45, 46]; (2) Massachusetts Institute of Technology (MIT) group under Prof.'s Jacob White and Dimitri Antoniadis [17–19]; and (3) the University of Maryland group under Prof.'s Neil Goldsman and Isaak Mayergoyz [20, 22–26, 32–34, 40–44, 47–51].

The spherical-harmonic expansion in space-independent formulation [17, 20–30, 40, 47–50] has been used to study effects like impact ionization [24–27, 29, 37]. Soon many many devices were simulated: one-dimensional n^+nn^+ device [17,19,32,33,42,48]; one-dimensional bipolar junction transistors (BJT) [31, 34, 36, 37, 44]; and two-dimensional MOSFET's [39, 41–46, 50, 51].

Some of the space-dependent simulations were a post-processing to the drift-diffusion or hydrodynamic simulation [32,36,37,39,41]; some simulations were self-consistent with the Poisson equation [17–19,33]; and a few simulation went beyond the Poisson equation and included the hole-continuity equation self-consistently [34, 42, 43, 51]. A hybrid simulator which combines efficient hydrodynamic with the spherical-harmonic approach was proposed in [45].

In terms of band-structure, the initial simulations used a simple elliptical

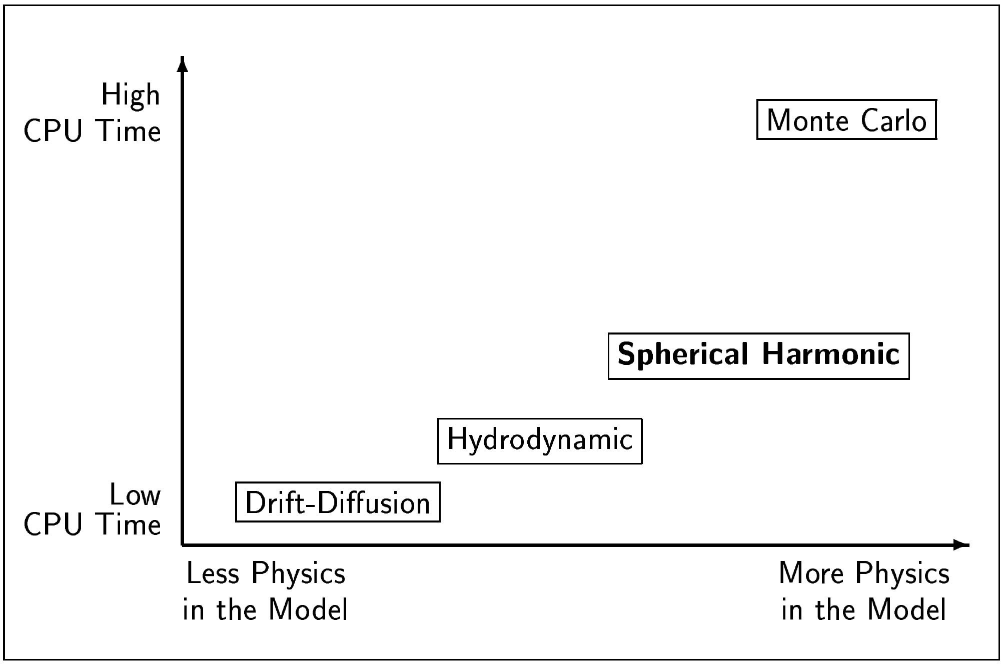

High
CPU Time
Low
CPU Time
Monte Carlo
Spherical Harmonic
Hydrodynamic
Drift-Diffusion
Less Physics
in the Model
More Physics
in the Model

Figure 1.2: A schematic representation of the physical rigor of the various transport models and the computational burden associated with them. Spherical-harmonic approach incorporates more physics in its formulation, but takes significantly less computation time.

parabolic band-structure [20, 32, 33, 40, 47, 48]. Soon a spherical multi-band band-structure by Brunetti *et al.* [57], and improvements on it [58], gained popularity. Although this band-structure has four bands, space-dependent simulations have used only the lower two [21, 37, 39, 42, 43]. All four bands have been incorporated only in space-independent formulations [23, 25, 50]. Vecchi *et al.* [27–29] have proposed an interesting idea to incorporate full-band band-structure, similar to the one used in full-band Monte Carlo, into the spherical-harmonic approach.

1.6 Scope & Contributions of the Dissertation

This dissertation makes further progress in the spherical-harmonic techniques for solving the Boltzmann transport equation. A major salient feature is that the space-dependent Boltzmann equation is solved for energies up to 3.4eV by including all four bands of the multi-band band-structure of Brunetti *et al.* [57], and its improvement [58]. All four bands, so far, have been used only in space-independent formulations [23, 25, 50]. In space-dependent formulations, the most that has been incorporated are the lower two bands [21, 37, 39, 42, 43]. Modeling the higher bands is important because electrons in modern short-geometry devices routinely reach high energies. These energetic electrons are also a major reliability concern.

In this dissertation we transform the spherical-harmonic-expanded Boltzmann equation to a new variable, called pre-Maxwellian variable. It is shown here that if the Boltzmann equation is transformed to use this new variable then its numerical properties improve, making the numerical solution more robust. This new variable transformation (a) enhances the diagonal dominance of the discretized Boltzmann equation; (b) accounts for exponential variation of distribution in both energy and space; (c) allows usage of superior Poisson-equation solvers, like Mayergoyz' fixed-point technique [95–97]; (d) while retaining the linearity of the Boltzmann equation intact. The Boltzmann equation in each band is transformed to this new variable and discretized using the control-volume formulation [113, 114]. The discretized equations are solved using line SOR in space and energy alternatively [114]. The solution exhibited *numerical noise.* The cause of this noise was identified and removed by inclusion of acoustic phonons in inelastic approximation, instead of the usual elastic approximation. The inelastic acoustic phonons scattering term was formulated in a new self-adjoint form, and a new stable discretization was developed. The use of inelastic acoustic phonons, it may be pointed out, is not an original contribution; it has been used in the original Spherical-Harmonic paper of Prof. Neil Goldsman [20]. What *is* new is that (a) it was identified as the cure for the numerical noise, and (b) a new discretization was developed based on expressing the scattering term in a self-adjoint form.

This approach of solving the Boltzmann equation was tested on three devices: (a) a one-dimensional n^+nn^+ device structure; (b) a one-dimensional bipolar junction transistor (BJT); and (c) a two-dimensional metal-oxide-semiconductor field-effect transistor (MOSFET). For the n^+nn^+ and BJT the Boltzmann equation was solved self-consistently with the Poisson and hole-continuity equations. For the MOSFET the Boltzmann equation was solved as a post-processor to a drift-diffusion simulation.

The case of thermal equilibrium was examined. An analytical solution to the multi-band Boltzmann-Poisson system was developed and used to validate the code. It was shown that despite large electric fields in devices, equilibrium conditions prevail in the device. At thermal equilibrium, electron velocity is zero; this can be interpreted as *velocity undershoot* or *velocity damping*, since the velocity is less than what is demanded by the local electric field.

1.7 Organization of the Dissertation

The dissertation is organized as follows. Chapter 2 on page 11 derives the basic mathematical formulation of the spherical-harmonic-expansion approach. The multi-band band-structure is discussed in Chapter 3 on page 43. The multi-band spherical-harmonic Boltzmann equation is transformed to the new pre-Maxwellian variable and discretized in Chapter 4 on page 53. This discretization, along with the Poisson and hole-continuity equation, is solved for the n^+nn^+ device in Chapter 5 on page 81. A bipolar junction transistor (BJT) is simulated in Chapter 6 on page 107, and a MOSFET is simulated Chapter 7 on page 133. Finally the dissertation concludes in Chapter 8 on page 171.

To maintain lucidity, some topics are discussed in the appendices. Proof of current-conserving property of the Boltzmann equation is given in Appendix A on page 173. Analytical solution to the thermal-equilibrium Boltzmann equation is derived in Appendix B on page 177. The derivation of harmonic-mean discretization, and its similarities to other schemes, is presented in Appendix C on page 179. Following that, Appendix D on page 183 introduces the Poisson equation along with its discretization, and Appendix E on page 189 presents the discretization of the hole-continuity equation.

1.8 Chapter Summary

In this chapter the central ideas of the dissertation were introduced. An argument was made that device simulation is an important area for semiconductors. Then it was indicated that the present simulation approaches—drift-diffusion, hydrodynamic, and Monte Carlo—have deficiencies which the spherical-harmonic approach addresses. Following that, a brief review of the spherical harmonic literature was given. The scope and contributions of this dissertation were discussed, followed by a description of its organization.

CHAPTER 2

SPHERICAL HARMONIC EXPANSION OF BOLTZMANN TRANSPORT EQUATION: DERIVATIONS AND FUNDAMENTALS

2.1 Chapter Introduction

Electron transport in semiconductors can be described by the semi-classical Boltzmann transport equation. Numerical simulation of semiconductor devices by this equation is a formidable task because of many reasons: (a) the dependent variable, distribution function, is a function of seven dimensions, three in real space, three in wave-vector space, and one in time; this is referred to as the *the curse of dimensionality.* (b) The Boltzmann transport equation is a complicated integro-differential equation. Over the last few decades many methods have been developed for its solution [12]: Monte Carlo method [62–67, 69, 70, 92]; scattering-matrix approach [13, 91]; deterministic particle method [14]; cellular-automata approach [15]; and iterative-spectral method [16] amongst others. Spherical-harmonic-expansion approach is an efficient direct numerical solution technique to solve the Boltzmann transport equation [17–51]. This chapter is devoted to presenting the basic derivations pertaining to this method.

Since considerable literature on the spherical-harmonic approach has accumulated in the recent past [17–51], this chapter, therefore, tries to eschew from re-deriving the old equations—final results will be presented with the appropriate citations. Impact-ionization scattering, however, is treated in more detail, this is because our derivation is at slight variance with literature [24, 43, 51].

The chapter is organized as follows. First the Boltzmann transport equation is introduced in Section 2.2. An overview of the derivation and equations for the spherical-harmonic approach is presented in Section 2.3 on page 14. In Section 2.4 on page 19 are given the derivations pertaining to average quantities, like electron concentration and current; and some functions, like density of states. We then introduce the expressions for the collision terms in Section 2.5 on page 26. All terms are put together into a differential-difference equation in Section 2.6 on page 34, which, along with the boundary conditions of Section 2.7 on page 35, represents the boundary-value problem to be solved in subsequent chapters.

2.2 Boltzmann Transport Equation

Let us assume that the distribution of electrons in real space, $\mathbf{r}$, and wave-vector space, $\mathbf{k}$, at time t can be described by a *distribution function* $f = f(\mathbf{r}, \mathbf{k}, t)$. We want this distribution function to have the following physical interpretation: It is the probability of finding an electron at a electronic state at $(\mathbf{r}, \mathbf{k})$ at time t. The total number of electrons in a given volume, therefore, is the sum of probabilities over all possible states: $\sum f(\mathbf{r}, \mathbf{k})$. This, as will be shown in Section 2.4, can be expressed as an integral $1/4\pi^3 \int f(\mathbf{r}, \mathbf{k})\, d^3\mathbf{r}\, d^3\mathbf{k}$. The distribution function, it follows, is a dimensionless function which describes how the electrons are distributed in real and wave-vector space [74].

Distribution function is a complete description of all electrons in a device—all quantities of interest in a device can be computed from the distribution function. The distribution function is governed by the celebrated *Boltzmann transport equation*. It is named after the renowned scientist Ludwig Boltzmann, shown in Fig. 2.1 on the facing page.

The Boltzmann transport equation can be derived as follows. Following the derivation by Tomizawa ([62], p. 171) we write the Boltzmann transport equation as:

$$\frac{\partial f}{\partial t} = -\left(\frac{\partial f}{\partial t}\right)_{\text{drift}} + \left[\frac{\partial f}{\partial t}\right]_{\text{collision}} \tag{2.1}$$

The drift term is the drift in both real and wave-vector space:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{drift}} = \frac{d\mathbf{r}}{dt} \cdot \nabla_r f + \frac{d\mathbf{k}}{dt} \cdot \nabla_k f \tag{2.2}$$

Using quantum mechanics we can write the *group velocity*, $\mathbf{u}_g$, in terms of the band structure.

$$\frac{d\mathbf{r}}{dt} = \mathbf{u}_g(\mathbf{k}) = \frac{1}{\hbar}\nabla_k \varepsilon \tag{2.3}$$

Figure 2.1: Ludwig Boltzmann (1844–1906). After his doctorate in 1866 he studied under many teachers like Josef Stefan, Bunsen, Kirchoff, and Helmholtz. It is said that he became depressed by arguments with his colleague and scientific opponent W. Ostwald and attempted suicide. He obtained the Maxwell-Boltzmann distribution in 1871. In 1884 he derived Josef Stefan's empirical T^4 law for black-body radiation from thermodynamic principles. In the 1890's he derived the Second Law of Thermodynamics from the principles of mechanics. In 1904 he visited Berkeley and Stanford. He committed suicide before the new discoveries concerning radiation, which he learned in his American trip, would prove his theories. After [11].

where the reduced Planck's constant, $\hbar$, is Planck's constant divided by 2π; and $\varepsilon = \varepsilon(\mathbf{k})$ is the energy of the electron, whose dependence on wave-vector, $\mathbf{k}$, comes from the knowledge of semiconductor band-structure. The momentum of electrons, by quantum mechanics, is $\mathbf{p} = \hbar\mathbf{k}$, and the rate of change of momentum, by classical mechanics, is the force acting on it:

$$\frac{d\mathbf{p}}{dt} = \hbar\frac{d\mathbf{k}}{dt} = -q\mathbf{E} \tag{2.4}$$

where the q is magnitude of charge on an electron, and $\mathbf{E}$ is the electric field. Substituting Eq. (2.4) and Eq. (2.3) in Eq. (2.2) and Eq. (2.1) we get the well-known Boltzmann transport equation:

$$\boxed{\frac{\partial f}{\partial t} + \frac{1}{\hbar}\nabla_k\varepsilon \cdot \nabla_r f - \frac{1}{\hbar}q\mathbf{E} \cdot \nabla_k f = \left[\frac{\partial f}{\partial t}\right]_{\text{collision}}} \tag{2.5}$$

where the nabla, ∇, operators have the following definitions.

$$\nabla_r \cdot \mathbf{A} \triangleq \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{2.6}$$

$$\nabla_k \cdot \mathbf{A} \triangleq \frac{\partial A_x}{\partial k_x} + \frac{\partial A_y}{\partial k_y} + \frac{\partial A_z}{\partial k_z} \tag{2.7}$$

Boltzmann transport equation, Eq. (2.5), describes electron transport in semiconductors in a semi-classical fashion. The electron motion is modeled by a combination of classical and quantum mechanics. The equation is essentially a statement of conservation in real space and wave-vector space. The Boltzmann equation is an integro-differential equation which has a high dimensionality—it has seven dimensions: three in real space; three in wave-vector space; and one in time. Because of this curse of dimensionality, a direct solution is rather difficult. Using spherical harmonics this dimensionality in wave-vector space is reduced and this integro-differential equation is transformed into a differential-difference equation, which is more amenable to a numerical solution.

2.3 Spherical Harmonic Expansion of the Boltzmann Equation

Distribution function can be expected to have some spherical symmetry in wave-vector space due to the spherical symmetry of the band-structure and the randomizing nature of the scattering processes [17]. In the limiting case, the Maxwellian distribution at thermal equilibrium is spherically symmetric. This *a priori* knowledge of symmetry can be exploited by expanding the distribution in wave-vector space using spherical-harmonic functions.

Spherical-harmonic-expansion approach can also be interpreted as a *spectral* method in azimuthal- and polar-angle direction, and the usual finite-difference in radius (energy) and space. This is similar to the approach in the discipline of computational fluid dynamics (CFD) [114,115]. Spectral methods in numerical solution of partial differential equations are well-known for their high accuracies arising from low discretization error [115].

Spherical harmonics are also used in other areas of science where there is spherical symmetry. In meteorological weather simulation, for example, spherical harmonics are used to solve for quantities at or near earth's surface. The earth provides natural spherical symmetry.

The spherical-harmonic expansion of the Boltzmann Transport equation has been treated in detail in literature: Hennacy [41,49] and the Italian group [35] originally derived many of these important results. Readers may also refer to [50, 51]. Therefore there is little need to re-derive these equations. I will only give a brief sketch of the derivations; an interested reader is referred to the original publications. The derivation of the spherical-harmonic expansion of the Boltzmann equation can be summarized in the following steps:

1. Expand the distribution in spherical harmonics (Eq. (2.8))
2. Substitute the expansion in the Boltzmann transport equation (Eq. (2.14))
3. Use orthogonality of spherical harmonic functions (Eq. (2.13)) to derive equations in the coefficients of the expansion
4. Convert the wave-vector magnitude, k, to energy, ε using the dispersion relation for the band-structure (Eq. (2.16))
5. Change the independent variable from energy to Hamiltonian (Eq. (2.19) and Eq. (2.20))
6. Truncate the infinite expansion to first-order (Eq. (2.21))

Referring to Fig. 2.2 on the next page we can express the wave-vector, $\mathbf{k} = (k_x, k_y, k_z)$, in terms of its spherical components (k, θ, ϕ). Where $k = |\mathbf{k}| =$ magnitude of electron wave vector, $\theta =$ polar angle, and $\phi =$ azimuthal angle. We can expand the distribution in terms of spherical harmonics as

$$\boxed{f(\mathbf{r}, \mathbf{k}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m(\mathbf{r}, k) Y_l^m(\theta, \phi) \qquad m \in [-l, l] \text{ and } l \in [0, \infty)} \tag{2.8}$$

The magnitude k has a physical significance—for spherical band-structure all points on a sphere have the same energy.

The functions Y_l^m are the spherical-harmonics, integers $m \in [-l, l]$, and $l \in [0, \infty)$. For some lower values of l and m these functions have the following

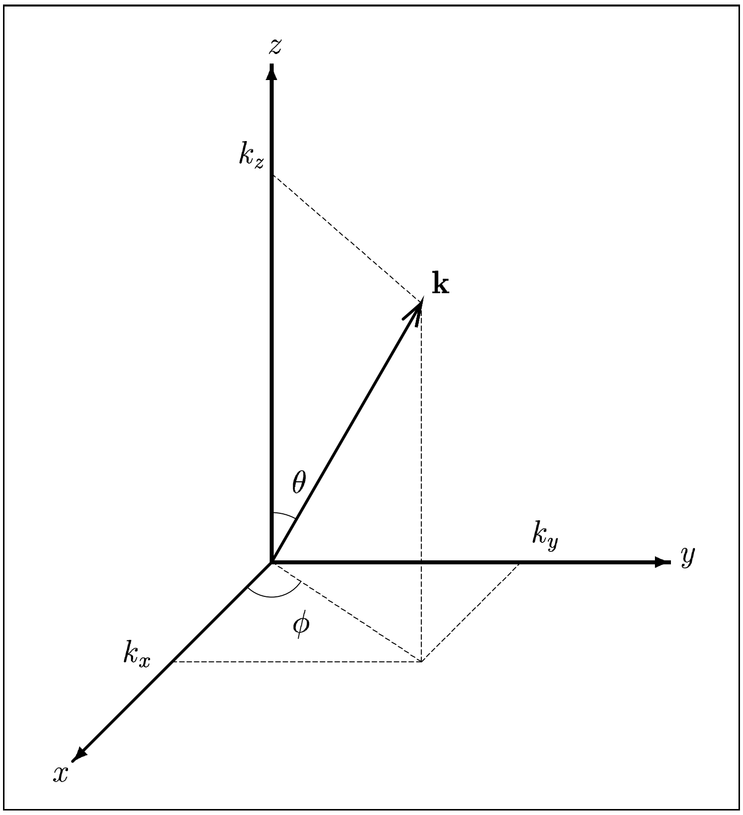


Figure 2.2: Coordinate systems used in the spherical-harmonic expansion. The coordinates of wave-vector $\mathbf{k}$ in the spherical system are the magnitude k, the polar angle θ, and the azimuthal angle ϕ. In the Cartesian system the wave-vector $\mathbf{k}$ will have familiar Cartesian components: k_x, k_y, and k_z.

form [71]:

$$Y_0^0(\theta,\phi) = \frac{1}{\sqrt{4\pi}} \tag{2.9}$$

$$Y_1^1(\theta,\phi) = -\sqrt{\frac{3}{8\pi}}\ \sin\theta\ e^{j\phi} \tag{2.10}$$

$$Y_1^0(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\ \cos\theta \tag{2.11}$$

$$Y_1^{-1}(\theta,\phi) = +\sqrt{\frac{3}{8\pi}}\ \sin\theta\ e^{-j\phi} \tag{2.12}$$

where $j = \sqrt{-1}$. Orthogonality of spherical-harmonic functions is an important property:

$$\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} Y_{l_1}^{m_1 *}(\theta,\phi) Y_{l_2}^{m_2}(\theta,\phi) \sin\theta\, d\theta\, d\phi = \delta_{l_1,l_2}\delta_{m_1,m_2} \tag{2.13}$$

Substituting the expansion of $f(\mathbf{r}, \mathbf{k})$ from Eq. (2.8) into the Boltzmann equation Eq. (2.5) we get

$$\left(\frac{\partial}{\partial t} + \frac{1}{\hbar}\nabla_k\varepsilon \cdot \nabla_r - \frac{1}{\hbar}q\mathbf{E} \cdot \nabla_k\right) \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m(\mathbf{r}, k) Y_l^m(\theta, \phi) = \left[\frac{\partial \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m(\mathbf{r}, k) Y_l^m(\theta, \phi)}{\partial t}\right]_{\text{collision}} \tag{2.14}$$

By projecting this on the basis functions

$$\int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} d\theta\, d\phi\ \sin\theta\, Y_l^{m*}(\theta, \phi) \left(\frac{\partial}{\partial t} + \frac{1}{\hbar}\nabla_k\varepsilon \cdot \nabla_r - \frac{1}{\hbar}q\mathbf{E} \cdot \nabla_k - \left[\frac{\partial}{\partial t}\right]_{\text{collision}}\right) \cdot \sum_{l'=0}^{\infty} \sum_{m=-l'}^{l} f_{l'}^{m'}(\mathbf{r}, k) Y_{l'}^{m'}(\theta, \phi) = 0 \tag{2.15}$$

we derive a set of coupled partial differential equations in the coefficients f_l^m.

We converted the independent variable from $\mathbf{k}$ to k, which can be further converted to the variable energy ε by the dispersion relation. By using the band-structure information we can write a dispersion relation $\gamma(\varepsilon)$ in terms of an effective mass m^*:

$$\boxed{\gamma(\varepsilon) = \frac{\hbar^2 k^2}{2m^*}} \tag{2.16}$$

The function $\gamma(\varepsilon)$ can be an arbitrary function, though it is usually of the form $\gamma(\varepsilon) = \varepsilon + \alpha\varepsilon^2$, or simply $\gamma(\varepsilon) = \varepsilon$.

In fact, the equations take on a simpler form if we go a step further and convert the independent variable from ε to the Hamiltonian, H [35, 36]. For an electron (negative charge), we can transform:

$$(\mathbf{r}, \varepsilon) \to (\mathbf{r}, H) \tag{2.17}$$

$$(x, y, z, \varepsilon) \to (x, y, z, H) \tag{2.18}$$

$$H = \varepsilon - q\phi(\mathbf{r}) \tag{2.19}$$

We can define new variable F_l^m when using H.

$$F_l^m(\mathbf{r}, H) \equiv f_l^m(\mathbf{r}, H + q\phi(\mathbf{r})) = f_l^m(\mathbf{r}, \varepsilon) \tag{2.20}$$

First-Order Truncation

Eq. (2.8) has infinite number of terms. It has been shown in [44] that a first-order expansion is accurate by comparing it with the exact Monte Carlo solution. Computations in [21] for second-order expansion compare well with those from first-order; this provides further justification for a first-order truncation. This should not be surprising since the spherical-harmonic approach can be interpreted as a spectral method. Spectral methods are accurate with lower-order expansion, in fact that is why they are used in the first place. The error with spectral methods decreases exponentially with increasing the number of terms of the expansion [115]. Motivated by this, we can truncate the infinite series in Eq. (2.8) to the *first order*: That is, we set $F_l^m = 0$, for $l \geq 2$. The distribution then looks like

$$\begin{aligned} f(\mathbf{r}, \mathbf{k}) &= \sum_{l=0,1} \sum_{m=-l}^{l} F_l^m(\mathbf{r}, H) Y_l^m(\theta, \phi) \\ &= F_0^0 Y_0^0 + \sum_{m=-1,0,1} F_1^m Y_1^m \qquad (2.21) \\ &= F_0^0 Y_0^0 + F_1^{-1} Y_0^{-1} + F_1^0 Y_0^0 + F_1^1 Y_0^1 \end{aligned}$$

Using this truncation, ignoring time-dependence and considering only one space-dimension, we can derive

$$\frac{u_g}{3} \left\{ \frac{\partial}{\partial x} \left(\tau u_g \frac{\partial F_0^0}{\partial x} \right) - qE \frac{\gamma'}{\gamma} \left(\tau u_g \frac{\partial F_0^0}{\partial x} \right) \right\} + \left[\frac{\partial F_0^0}{\partial t} \right]_{\text{collision}} = 0 \qquad (2.22)$$

where $\tau(\varepsilon)$ is the reciprocal of the scattering rate; $\gamma' = d\gamma(\varepsilon)/d\varepsilon$ is the derivative of the dispersion relation; and u_g is the group velocity, to be discussed in Eq. (2.55). We have, in (2.22), abused notation slightly by retaining the symbols for u_g, τ, γ, even though the independent variable has been changed from ε to H. This equation can be re-arranged in a self-adjoint form:

$$\frac{u_g}{3\gamma} \frac{\partial}{\partial x} \left(\tau u_g \gamma \frac{\partial F_0^0}{\partial x} \right) + \left[\frac{\partial F_0^0}{\partial t} \right]_{\text{collision}} = 0 \qquad (2.23)$$

By noticing $u_g = (\sqrt{2/m^*})(\sqrt{\gamma}/\gamma')$ and $g = g(\varepsilon) = c_g \sqrt{\gamma} \gamma'$ we see that $g \propto \gamma/u_g$; which allows us to write [31, 35–37, 39]

$$\boxed{\frac{1}{3g} \frac{\partial}{\partial x} \left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x} \right) + \left[\frac{\partial F_0^0}{\partial t} \right]_{\text{collision}} = 0} \qquad (2.24)$$

We will refer to Eq. (2.24) as the *spherical-harmonic Boltzmann equation.*

The spherical-harmonic expansion has some interesting consequences:

1. Original Boltzmann equation, ignoring time-dependence, had six dimensions: three in real-space, three in wave-vector space. The spherical-harmonic Boltzmann equation has four: three in real-space, one in Hamiltonian.
2. Original Boltzmann equation was a first-order integro-differential equation, the spherical-harmonic Boltzmann equation is second-order differential equation. This is makes it easier to specify boundary conditions in a boundary-value problem.
3. The lower-order expansion coefficients have physical significance by themselves. f_0^0 is the isotropic symmetric part of the distribution—physically it corresponds to the energy distribution function, it is a measure of how electrons are distributed in energy. $f_1^{\pm 1}$ and f_1^0 are the asymmetric part of the distribution—physically they describes electron flow (current).

To proceed further we need expressions for the scattering terms. These are derived in Section 2.5 on page 26. But before that we need to express the average quantities in terms of spherical-harmonic expansion.

2.4 Average Quantities

2.4.1 Basics

For a function $Q(\mathbf{k})$ we can compute its sum over all electrons in a volume Ω as

$$< NQ >= \sum_{\text{all } e^- \text{ states in } k\text{-space}} Q(\mathbf{k}) f(\mathbf{k}) \tag{2.25}$$

where N is the total number of electrons in in the volume. The average, or moment of distribution, of the quantity $Q(\mathbf{k})$ for all electrons in the volume is defined as:

$$< Q > \triangleq \frac{< NQ >}{N} \tag{2.26}$$

2.4.2 Summation Over k-Space

Solution of Schrödinger equation for electrons in the periodic potential of a semiconductor gives [74]

$$\Delta k_x = \frac{2\pi}{L_x} \tag{2.27}$$

where L_x is the length of the sample in x-direction. Using similar expressions in the other spatial directions we can write

$$\begin{aligned} \Delta k_x \, \Delta k_y \, \Delta k_z &= \text{Volume of one } k\text{-space state} \\ &= \frac{(2\pi)^3}{L_x L_y L_z} \\ &= \frac{(2\pi)^3}{\Omega} \end{aligned} \tag{2.28}$$

Where $\Omega = L_x L_y L_z =$ is the volume of the crystal; volume Ω can be thought of as a conceptual box whose dimensions are large compared to the electron wavelength, but small as compared to the size of a device [74]. For every $k-$space element there can be two electrons—one with up-spin and the other with down-spin. Including this factor of two, we can write

$$\frac{\text{Number of Electron States}}{k\text{-space Volume}} = \frac{\Omega}{(2\pi)^3} \times 2 \tag{2.29}$$

If a function $A(\mathbf{k}) = A(k_x, k_y, k_z)$ is a function of wave-vector. Sum of $A(\mathbf{k})$ over all electronic states is

$$\begin{aligned} \sum_{\text{all } e^- \text{ states}} A(\mathbf{k}) &= 2 \sum_{\text{all } \mathbf{k}} A(\mathbf{k}) \\ &= \frac{1}{\Delta k_x \Delta k_y \Delta k_z} \sum_{\text{all}\mathbf{k}} A(\mathbf{k}) \; \Delta k_x \Delta k_y \Delta k_z \\ &= 2 \, \frac{\Omega}{(2\pi)^3} \int_{\text{all } \mathbf{k}} A(\mathbf{k}) \, d^3\mathbf{k} \end{aligned} \tag{2.30}$$

We have derived a general rule to convert a sum to an integral.

$$\boxed{\sum_{\mathbf{k}} \quad \longrightarrow \quad 2 \, \frac{\Omega}{(2\pi)^3} \int_{\mathbf{k}} \quad d^3\mathbf{k}} \tag{2.31}$$

2.4.3 Integration Over k-Space

Let us say there is a function $B = B(k)$ which depends only on the magnitude of wave-vector, k. Integration of such a function over the k-space takes on a special frequently-used form, which we derive in this sub-section. In spherical coordinates the differential k-space element is

$$d^3\mathbf{k} = (dk) \, (k \, d\theta) \, (k \sin\theta \, d\phi) = k^2 \sin\theta \, dk \, d\theta \, d\phi \tag{2.32}$$

Integrating over the entire k-space

$$\begin{aligned}\int_{\mathbf{k}} B(k)\, d^3\mathbf{k} &= \int_{k=0}^{k_{\max}} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} B(k)\ k^2 \sin\theta\, dk\, d\theta\, d\phi \\ &= \int_{k=0}^{k_{\max}} B(k)\ k^2\, dk \int_{\theta=0}^{\pi} \sin\theta\, d\theta \int_{\phi=0}^{2\pi} d\phi \\ &= 4\pi \int_k B(k)\ k^2\, dk\end{aligned} \tag{2.33}$$

From the dispersion relation of Eq. (2.16), $\gamma(\varepsilon) = \hbar^2 k^2/2m^*$, we can derive

$$\gamma'(\varepsilon)\, d\varepsilon = \frac{\hbar^2 k}{m^*}\, dk \tag{2.34}$$

$$k\, dk = \frac{m^*}{\hbar^2}\ \gamma'(\varepsilon)\, d\varepsilon \tag{2.35}$$

$$k = \frac{\sqrt{2m^*\gamma(\varepsilon)}}{\hbar} \tag{2.36}$$

We can then introduce a new function g

$$\begin{aligned}k^2\, dk &= \frac{\sqrt{2}(m^*)^{3/2}}{\hbar^3}\gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)}\ d\varepsilon \\ &= \frac{g(\varepsilon)}{4\pi}\, d\varepsilon\end{aligned} \tag{2.37}$$

Which is defined as:

$$\boxed{g(\varepsilon) \triangleq \frac{4\pi\sqrt{2}(m^*)^{3/2}}{\hbar^3}\ \gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)}} \tag{2.38}$$

For the sake of brevity we would often define a constant $c_g = 4\pi\sqrt{2}(m^*)^{3/2}/\hbar^3$.

The complete the integration in Eq. (2.33) we change variable of integration from $k \to \varepsilon$. The function B in new variable is

$$B(k) = B\left(\frac{\sqrt{2m^*\gamma(\varepsilon)}}{\hbar}\right) = \mathcal{B}(\varepsilon) \tag{2.39}$$

Substituting Eq. (2.39) and Eq. (2.37) in Eq. (2.33)

$$\begin{aligned}\int_{\mathbf{k}} B(k)\, d^3\mathbf{k} &= 4\pi \int_k B(k)\ k^2\, dk \\ &= 4\pi \int_\varepsilon \mathcal{B}(\varepsilon)\frac{\sqrt{2}(m^*)^{3/2}}{\hbar^3}\gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)}\ d\varepsilon \\ &= \int_\varepsilon \mathcal{B}(\varepsilon) g(\varepsilon)\ d\varepsilon\end{aligned} \tag{2.40}$$

This translates to a simple rule: Integration over entire k-space of a function which depends only on wave-vector magnitude can be done instead by multiplying it by $g(\varepsilon)$ and integrating over energy.

$$\boxed{\int_{\mathbf{k}} B(k)\ d^3\mathbf{k} \qquad \longrightarrow \qquad \int_{\varepsilon} \mathcal{B}(\varepsilon) g(\varepsilon)\ d\varepsilon} \tag{2.41}$$

2.4.4 Electron Concentration

Electron concentration at a point in space can be computed by setting $Q = 1$ in Eq. (2.25), and applying Eq. (2.31) and Eq. (2.32).

$$\begin{aligned} n &= \frac{\text{Total Number of Electrons}}{\text{Volume}} \\ &= \frac{N}{\Omega} \\ &= \frac{1}{\Omega} \sum_{\text{all } e^- \text{ states}} f(\mathbf{k}) \\ &= \frac{2}{(2\pi)^3} \int_{\text{all } \mathbf{k}} f(\mathbf{k})\ d^3\mathbf{k} \\ &= \frac{2}{(2\pi)^3} \int_{k=0}^{k_{\max}} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} f(\mathbf{k})\ k^2 \sin\theta\, dk\, d\theta\, d\phi \end{aligned} \tag{2.42}$$

$k_{\max}$ can be ∞ or the maximum of a band. Substituting the expansion $f(\mathbf{k}) = \sum_l \sum_m f_l^m(k) Y_l^m(\theta, \phi)$ from Eq. (2.8)

$$\begin{aligned} n &= \frac{2}{(2\pi)^3} \int_k \int_\theta \int_\phi f(\mathbf{k})\ k^2 \sin\theta\, dk\, d\theta\, d\phi \\ &= \frac{2}{(2\pi)^3} \int_k \int_\theta \int_\phi \sum_l \sum_m f_l^m(k) Y_l^m(\theta, \phi) \quad k^2 \sin\theta\, dk\, d\theta\, d\phi \\ &= \frac{2}{(2\pi)^3} \sum_l \sum_m \int_k f_l^m(k)\ k^2\, dk \int_\theta \int_\phi Y_l^m(\theta, \phi) \sin\theta\, d\theta\, d\phi \end{aligned} \tag{2.43}$$

From Eq. (2.9), $Y_0^0 = Y_0^{0*} = 1/\sqrt{4\pi}$. Then $Y_0^{0*}\sqrt{4\pi}$ is simply unity. We can then apply the orthogonality of spherical harmonics, Eq. (2.13), as follows:

$$\begin{aligned}
n &= \frac{2}{(2\pi)^3}\sum_l\sum_m \int_k f_l^m(k)\, k^2\, dk \int_\theta\int_\phi Y_l^m(\theta,\phi)\frac{\sqrt{4\pi}}{\sqrt{4\pi}}\sin\theta\, d\theta\, d\phi \\
&= \frac{2}{(2\pi)^3}\sqrt{4\pi}\,\sum_l\sum_m \int_k f_l^m(k)\, k^2\, dk \int_\theta\int_\phi Y_l^m(\theta,\phi)Y_0^{0*}\sin\theta\, d\theta\, d\phi \\
&= \frac{2}{(2\pi)^3}\sqrt{4\pi}\,\sum_l\sum_m \int_k f_l^m(k)\, k^2\, dk\ \cdot\ \delta_{l,0}\delta_{m,0} \qquad (2.44)\\
&= \frac{2}{(2\pi)^3}\sqrt{4\pi}\,\int_k f_0^0(k)\, k^2\, dk \\
&= \frac{2}{(2\pi)^3}\sqrt{4\pi}\,\int_\varepsilon f_0^0(\varepsilon)\left(\frac{\sqrt{2}(m^*)^{3/2}}{\hbar^3}\gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)}\right)\, d\varepsilon
\end{aligned}$$

to contain the burgeoning number of symbols we have flouted mathematical rigor slightly by setting $f_0^0(k) = f_0^0(\varepsilon)$. Re-arranging the above equation we get:

$$\boxed{n = \frac{(m^*)^{3/2}}{\pi^2\sqrt{2\pi}\ \hbar^3}\int_\varepsilon f_0^0(\varepsilon)\gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)}\, d\varepsilon} \qquad (2.45)$$

To put this equation in familiar terms, we can write it as

$$\begin{aligned}
n &= \int_\varepsilon \left(f_0^0(\varepsilon)\frac{1}{\sqrt{4\pi}}\right)\ h(\varepsilon)\, d\varepsilon \\
&= \int_\varepsilon \left(f_0^0(\varepsilon)Y_0^0\right)\ h(\varepsilon)\, d\varepsilon
\end{aligned} \qquad (2.46)$$

The function $h(\varepsilon)$ is the density of states (DOS) [76]. It is proportional to the function $g(\varepsilon)$ from Eq. (2.38)

$$h(\varepsilon) = \frac{1}{4\pi^3}g(\varepsilon) \qquad (2.47)$$

2.4.5 Electron Current and Velocity

To derive an expression for electron current density, or electron velocity, we set $Q = \mathbf{u}_g$, where $\mathbf{u}_g$ is the group velocity of electrons, in Eq. (2.25):

$$\begin{aligned}
< N\mathbf{u}_g > &= \sum_{\text{all } e^- \text{ states}} \mathbf{u}_g(\mathbf{k})f(\mathbf{k}) \\
&= 2\ \frac{\Omega}{(2\pi)^3}\int_{\mathbf{k}} \mathbf{u}_g(\mathbf{k})f(\mathbf{k})d^3\mathbf{k}
\end{aligned} \qquad (2.48)$$

Basic definition of Current density is carrier density times velocity.

$$\mathbf{J} \triangleq < n\mathbf{u}_g >= \frac{< N\mathbf{u}_g >}{\Omega} \tag{2.49}$$

and average velocity is defined by

$$< \mathbf{u}_g > \triangleq \frac{\mathbf{J}}{n} \tag{2.50}$$

Few things need to be derived to proceed further. From the dispersion relation Eq. (2.16) we write

$$\gamma\left(\varepsilon(\mathbf{k})\right) = \frac{\hbar^2 k^2}{2m^*} = \frac{\hbar^2}{2m^*}\left(k_x^2 + k_y^2 + k_z^2\right) \tag{2.51}$$

The chain rule of differentiation,

$$\frac{\partial \gamma}{\partial k_x} = \frac{\partial \gamma}{\partial \varepsilon} \cdot \frac{\partial \varepsilon}{\partial k_x} \tag{2.52}$$

can be re-arranged to derive

$$\begin{aligned} \frac{\partial \varepsilon}{\partial k_x} &= \frac{1}{\gamma'(\varepsilon)} \frac{\partial \gamma}{\partial k_x} \\ &= \frac{1}{\gamma'(\varepsilon)} \frac{\hbar^2}{2m^*} \frac{\partial}{\partial k_x}\left(k_x^2 + k_y^2 + k_z^2\right) \\ &= \frac{1}{\gamma'(\varepsilon)} \frac{\hbar^2}{2m^*}\left(2k_x\right) \\ &= \frac{1}{\gamma'(\varepsilon)} \frac{\hbar^2}{m^*} k_x \end{aligned} \tag{2.53}$$

From quantum mechanics the group velocity of electrons, as stated in Eq. (2.3), can be written in terms of the band structure:

$$\begin{aligned} \mathbf{u}_g(\mathbf{k}) &= \frac{1}{\hbar} \nabla_k \varepsilon(\mathbf{k}) \\ &= \frac{1}{\hbar}\left(\frac{\partial \varepsilon}{\partial k_x}\hat{\mathbf{i}} + \frac{\partial \varepsilon}{\partial k_y}\hat{\mathbf{j}} + \frac{\partial \varepsilon}{\partial k_z}\hat{\mathbf{k}}\right) \\ &= \frac{1}{\gamma'(\varepsilon)} \frac{\hbar}{m^*}\left(k_x\hat{\mathbf{i}} + k_y\hat{\mathbf{j}} + k_z\hat{\mathbf{k}}\right) \\ &= \frac{\hbar}{m^*} \frac{\mathbf{k}}{\gamma'(\varepsilon)} \end{aligned} \tag{2.54}$$

The symbols $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$, are unit vectors in x, y, z direction respectively. Magnitude of $\mathbf{u}_g(\mathbf{k})$ is

$$\begin{aligned} u_g &= \frac{\hbar}{m^*}\frac{k}{\gamma'(\varepsilon)} \\ &= \frac{\hbar}{m^*}\frac{1}{\gamma'(\varepsilon)}\left(\frac{\sqrt{2m^*\gamma(\varepsilon)}}{\hbar}\right) \\ &= \sqrt{\frac{2}{m^*}}\,\frac{\sqrt{\gamma(\varepsilon)}}{\gamma'(\varepsilon)} \end{aligned} \tag{2.55}$$

Let us consider the x components. From Eq. (2.55) and Fig. 2.2 on page 16

$$u_{g\,x} = u_g \sin\theta\,\cos\phi = \sqrt{\frac{2}{m^*}}\,\frac{\sqrt{\gamma(\varepsilon)}}{\gamma'(\varepsilon)}\,\sin\theta\,\cos\phi \tag{2.56}$$

Substituting in Eq. (2.49), the x component of current density is

$$\begin{aligned} J_x &=< n\ u_{g\,x} >= \frac{< N u_{g\,x} >}{\Omega} \\ &= \frac{2}{(2\pi)^3}\int_{\mathbf{k}} u_{g\,x}(k) f(\mathbf{k}) d^3\mathbf{k} \\ &= \frac{2}{(2\pi)^3}\int_{k=0}^{k_{\max}}\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}(u_g\ \sin\theta\ \cos\phi)\cdot \\ &\cdot\left(\sum_l\sum_m f_l^m(k) Y_l^m(\theta,\phi)\right) k^2\sin\theta\, dk\, d\theta\, d\phi \end{aligned} \tag{2.57}$$

Using the formulas for spherical harmonics, Eq. (2.10) and Eq. (2.12)

$$\sin\theta\ \cos\phi = \sqrt{\frac{2\pi}{3}}\left(Y_1^{-1} - Y_1^1\right) \tag{2.58}$$

we can simplify Eq. (2.57) by using orthogonality from Eq. (2.13)

$$\begin{aligned} J_x &= \frac{2}{(2\pi)^3}\int_{k=0}^{k_{\max}} u_g k^2\ \sqrt{\frac{2\pi}{3}} \\ &\cdot\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}\left(Y_1^{-1} - Y_1^1\right)\cdot\left(\sum_l\sum_m f_l^m(k) Y_l^m(\theta,\phi)\right)\sin\theta\, d\theta\, d\phi\ \ dk \\ &= \frac{2}{(2\pi)^3}\sqrt{\frac{2\pi}{3}}\int_{k=0}^{k_{\max}} u_g\cdot\left(f_1^{-1}(k) - f_1^1(k)\right) k^2\ dk \\ &= \frac{2}{(2\pi)^3}\sqrt{\frac{2\pi}{3}}\int_{\varepsilon} u_g\cdot\left(f_1^{-1}(\varepsilon) - f_1^1(\varepsilon)\right)\frac{g(\varepsilon)}{4\pi}\ d\varepsilon \end{aligned} \tag{2.59}$$

Now we invoke the transformation of variables from energy to Hamiltonian, as derived in Eq. (2.19) and Eq. (2.20).

$$f_1^1(\varepsilon) \longrightarrow F_1^1(H) \qquad \text{and} \qquad f_1^{-1}(\varepsilon) \longrightarrow F_1^{-1}(H) \tag{2.60}$$

For the Hamiltonian variable, simple expressions have been derived in [41, 49] and [50, 51].

$$F_1^1(H) = \frac{u_g\tau}{\sqrt{6}}\frac{\partial F_0^0}{\partial x} = -F_1^{-1}(H) \tag{2.61}$$

$$F_1^{-1}(H) - F_1^1(H) = -\sqrt{\frac{2}{3}}\, u_g\tau\frac{\partial F_0^0}{\partial x} \tag{2.62}$$

which can be substituted in Eq. (2.59) to read

$$\boxed{J_x = -\frac{1}{24\pi^3\sqrt{\pi}}\int_H g u_g^2\tau\frac{\partial F_0^0}{\partial x}\, dH} \tag{2.63}$$

Sometimes for brevity we will use $c_j = 1/24\pi^3\sqrt{\pi}$. Average velocity in x direction is J_x divided by the electron concentration n (Eq. (2.49)):

$$u_{g\,x} = \frac{J_x}{n} \tag{2.64}$$

2.5 Collision Terms

The collision term in the Boltzmann equation for non-degenerate semiconductors is (p. 109 in [74])

$$\begin{aligned}
\left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{collision}} &= \sum_{\text{all } e^- \text{ states}} f(\mathbf{k}')S(\mathbf{k}',\mathbf{k}) - f(\mathbf{k})S(\mathbf{k},\mathbf{k}') \\
&= \sum_{\mathbf{k}\,\uparrow} f(\mathbf{k}')S(\mathbf{k}',\mathbf{k}) - f(\mathbf{k})S(\mathbf{k},\mathbf{k}') \\
&= \frac{\Omega}{(2\pi)^3}\int_{\mathbf{k}'} f(\mathbf{k}')S(\mathbf{k}',\mathbf{k}) - f(\mathbf{k})S(\mathbf{k},\mathbf{k}')\; d^3\mathbf{k}'
\end{aligned} \tag{2.65}$$

Transition rate $S(\mathbf{k},\mathbf{k}')$ = probability of an electron scattering from state $\mathbf{k}$ to state $\mathbf{k}'$. The vertical arrow under the summation sign indicates that the sum over final states includes only those with spin parallel to the incident electrons; this is because scattering does not flip the sign of the electron (p. 46 in [74], or p. 99 in [75]). Therefore, the usual factor of two while converting the summation to an integral is omitted.

From the quantum mechanical theory of scattering based on Fermi's golden rule [62] we get the transition rates that are used in Eq. (2.65) to evaluate the scattering expressions. This leads to expressions like the following for all scattering processes [21,23,32,33,41,43,47–51]:

$$\begin{aligned}\left[\frac{\partial f(\mathbf{k})}{\partial t}\right]^{\text{total}}_{\text{coll}} &= \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]^{\text{in}}_{\text{coll}} - \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]^{\text{out}}_{\text{coll}} \\ &= \left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{coll}} Y_0^0 + \sum_{l:l\neq 0}\sum_m \left[\frac{\partial f_l^m(\varepsilon)}{\partial t}\right]_{\text{coll}} Y_l^m \\ &= \left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{coll}} Y_0^0 - \sum_{l:l\neq 0}\sum_m \frac{f_l^m(\varepsilon)}{\tau_{l,\text{coll}}(\varepsilon)} Y_l^m\end{aligned} \tag{2.66}$$

We include all major scattering mechanisms for silicon in our work:

1. Optical Phonon Scattering
2. Inter-Band Optical Phonon Scattering
3. Inelastic Acoustic Phonon Scattering
4. Ionized Impurity Scattering
5. Impact Ionization Scattering

For all scattering mechanism, save for impact ionization, we present only the final results; references [21,23,32,33,41,43,47–51] give details of the derivation. The derivation of impact ionization closely parallels the derivation of [24,43,51], but there are slight differences in the treatment of the scattering-in term. For this reason, impact ionization is outlined in slightly more details.

2.5.1 Optical Phonon Scattering

The transition rate for optical phonon scattering is given by (p. 68 of Lundstrom [74], p. 47 of Tomizawa [62], or [17,66]):

$$S(\mathbf{k},\mathbf{k}') = \frac{\pi \mathcal{D}_{\text{opt}}}{\rho\omega_{\text{opt}}\Omega}(N_{\text{opt}} + \frac{1}{2} \mp \frac{1}{2})\delta(\mathbf{k}' - \mathbf{k} \mp \mathbf{q})\,\delta(\varepsilon(\mathbf{k}') - \varepsilon(\mathbf{k}) \mp \hbar\omega_{\text{opt}}) \tag{2.67}$$

where $\mathcal{D}_{\text{opt}}$ is the optical deformation potential; ω_{opt} is the frequency of optical phonons, and $\hbar\omega_{\text{opt}}$, therefore, is the energy of optical phonons; $\mathbf{q}$ is the wave-vector of optical phonons; and ρ the density of silicon. Upper sign is for absorption of phonons; lower sign is for emission of phonons. Number of optical phonons, N_{opt}, are given by the Bose-Einstein factor for optical phonons (p. 39 Lundstrom [74]) as $N_{\text{opt}} = 1/(\exp(\hbar\omega_{\text{opt}}/k_B T_L) - 1)$, where T_L is the lattice temperature, and k_B the Boltzmann constant.

The scattering expressions derived for Y_0^0 and Y_l^m terms in Eq. (2.66) read

$$\begin{aligned}
\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{opt}} &= c_{\text{opt}}\, g^- \left(f_0^0(\varepsilon-\hbar\omega_{\text{opt}}) - e^{\hbar\omega_{\text{opt}}/k_BT_L} f_0^0(\varepsilon)\right) + \\
&\quad c_{\text{opt}}\, g^+ \left(e^{\hbar\omega_{\text{opt}}/k_BT_L} f_0^0(\varepsilon+\hbar\omega_{\text{opt}}) - f_0^0(\varepsilon)\right) \qquad (2.68)\\
\left[\frac{\partial f_l^m(\varepsilon)}{\partial t}\right]_{\text{opt}} &= -\frac{f_l^m(\varepsilon)}{\tau_{\text{opt}}(\varepsilon)} \qquad \text{For } l \neq 0 \qquad (2.69)
\end{aligned}$$

The scattering rate is given as

$$\frac{1}{\tau_{\text{opt}}(\varepsilon)} = c_{\text{opt}}\, g^- e^{\hbar\omega_{\text{opt}}/k_BT_L} + c_{\text{opt}}\, g^+ \qquad (2.70)$$

where $g^\pm = g(\varepsilon \pm \hbar\omega_{\text{opt}})$; the constant c_{opt} is

$$c_{\text{opt}} = \frac{\mathcal{D}_{\text{opt}}^2}{8\pi^2\rho\omega_{\text{opt}}} N_{\text{opt}} = \frac{\mathcal{D}_{\text{opt}}^2}{8\pi^2\rho\omega_{\text{opt}}} \frac{1}{e^{\hbar\omega_{\text{opt}}/k_BT_L}-1} \qquad (2.71)$$

2.5.2 Inter-Band Optical Phonon Scattering

Inter-band scattering is similar to optical phonon scattering. Let us consider a band (ν) which is interacting with another band (ν') through inter-band scattering: Electrons can scatter out of the current band (ν) to another band ν'; or, scatter in to band (ν) to from band (ν'). The scattering term is evaluated in a manner similar to optical phonons [23, 50].

$$\begin{aligned}
\left[\frac{\partial f_0^{0(\nu)}(\varepsilon)}{\partial t}\right]_{\text{ib}}^{(\nu\leftrightarrow\nu')} &= c_{\text{ib}}\, g^{-(\nu')} \left(f_0^{0(\nu')}(\varepsilon-\hbar\omega_{\text{ib}}) - e^{\hbar\omega_{\text{ib}}/k_BT_L} f_0^{0(\nu)}(\varepsilon)\right) + \\
&\quad c_{\text{ib}}\, g^{+(\nu')} \left(e^{\hbar\omega_{\text{ib}}/k_BT_L} f_0^{0(\nu')}(\varepsilon+\hbar\omega_{\text{ib}}) - f_0^{0(\nu)}(\varepsilon)\right) \qquad (2.72)\\
\left[\frac{\partial f_l^{m(\nu)}(\varepsilon)}{\partial t}\right]_{\text{ib}}^{(\nu\leftrightarrow\nu')} &= -\frac{f_l^{m(\nu)}(\varepsilon)}{\tau_{\text{ib}}^{(\nu\leftrightarrow\nu')}(\varepsilon)} \qquad \text{For} \quad l \neq 0 \qquad (2.73)
\end{aligned}$$

And the inter-band scattering rate reads

$$\frac{1}{\tau_{\text{ib}}^{(\nu\leftrightarrow\nu')}(\varepsilon)} = c_{\text{ib}}\, g^{-(\nu')} e^{\hbar\omega_{\text{ib}}/k_BT_L} + c_{\text{ib}}\, g^{+(\nu')} \qquad (2.74)$$

where $g^{\pm(\nu')} = g^{(\nu')}(\varepsilon \pm \hbar\omega_{\text{ib}})$; $f_0^{0(\nu')}(\varepsilon \pm \hbar\omega_{\text{ib}})$ are evaluated in the band (ν'); $f_0^{0(\nu)}(\varepsilon)$ is evaluated in band (ν). The constant c_{ib} reads

$$c_{\text{ib}} = \frac{\mathcal{D}_{\text{ib}}^2}{8\pi^2\rho\omega_{\text{ib}}} N_{\text{ib}} = \frac{\mathcal{D}_{\text{ib}}^2}{8\pi^2\rho\omega_{\text{ib}}} \frac{1}{e^{\hbar\omega_{\text{ib}}/k_BT_L}-1} \qquad (2.75)$$

Where the symbols $\mathcal{D}_{\text{ib}}$, $\hbar\omega_{\text{ib}}$, ω_{ib} have meanings similar to their counterparts in optical phonons.

2.5.3 Inelastic Acoustic Phonon Scattering

The transition rate for acoustic phonon scattering is (p. 44 in Tomizawa [62], or p. 66 in Lundstrom [74], or references [17, 66]):

$$S(\mathbf{k},\mathbf{k}') = \frac{\pi \mathcal{D}_{\mathrm{ac}}^2 q^2}{\rho \omega_q \Omega}(N_q + \frac{1}{2} \mp \frac{1}{2})\, \delta(\mathbf{k}' - \mathbf{k} \mp \mathbf{q})\, \delta(\varepsilon(\mathbf{k}') - \varepsilon(\mathbf{k}) \mp \hbar\omega_q) \qquad (2.76)$$

where $\mathcal{D}_{\mathrm{ac}}$ is the acoustic phonon deformation potential; q, not to be confused with the symbol for electronic charge, is the magnitude of acoustic phonon wave-vector $\mathbf{q}$; ω_q is the frequency of the acoustic phonons; ρ the density of silicon; and N_q is the number of acoustic phonons given by the Bose-Einstein relation. The upper sign in Eq. (2.76) is for absorption, and the lower one is for emission.

Acoustic phonons can be treated in either an elastic or inelastic approximation. Inelastic approximation was used in the initial spherical-harmonic work [20, 32, 33, 47, 48]; while in later work elastic approximation was used. Elastic approximation is sound from physical arguments [62, 74], but we use the inelastic approximation here for numerical reasons. Without inelastic acoustic phonons the distribution function displayed excessive numerical noise, as discussed in Section 5.8 on page 91.

Scattering term for inelastic acoustic phonon was derived by J. Kolodziejczak in 1967 [59]. Expression for the lowest-order term is

$$\begin{aligned} \left[\frac{\partial f_0^0}{\partial t}\right]_{\mathrm{ac}} &= c_{\mathrm{ac}}^{\mathrm{inelastic}} \sqrt{\gamma}\gamma'^2 \left\{ \left(1 + \frac{\gamma\gamma''}{\gamma'^2}\right) f_0^0 \right. \\ &\left. + \left(\frac{\gamma}{2\gamma'} + \left(1 + \frac{\gamma\gamma''}{\gamma'^2}\right) k_B T_L\right) \frac{\partial f_0^0}{\partial \varepsilon} + \frac{\gamma k_B T_L}{2\gamma'} \frac{\partial^2 f_0^0}{\partial \varepsilon^2} \right\} \end{aligned} \qquad (2.77)$$

where $f_0^0 = f_0^0(\varepsilon)$; and $c_{\mathrm{ac}}^{\mathrm{inelastic}} = (4\sqrt{2} m^{*5/2} \mathcal{D}_{\mathrm{ac}}^2)/(\pi\hbar^4\rho)$. Higher-order terms take on the familiar relaxation-type form:

$$\left[\frac{\partial f_l^m}{\partial t}\right]_{\mathrm{ac}} = -\frac{f_l^m}{\tau_{\mathrm{ac}}} \qquad \text{For} \qquad l \neq 0 \qquad (2.78)$$

Scattering rate is given by

$$\frac{1}{\tau_{\mathrm{ac}}(\varepsilon)} = \frac{\mathcal{D}_{\mathrm{ac}}^2\, k_B T_L}{(2\pi)^2 v_{\mathrm{sound}}^2\, \rho\, \hbar}\, g(\varepsilon) = c_{\mathrm{ac}}\, g(\varepsilon) \qquad (2.79)$$

where v_{sound} is the velocity of sound in silicon.

2.5.4 Ionized Impurity Scattering

Transition rate for ionized impurity scattering is (p. 35 Tomizawa [62]):

$$S(\mathbf{k},\mathbf{k}') = \frac{2\pi}{\hbar}\frac{N_I Z^2 q^4}{\Omega \epsilon_{\mathrm{si}}}\frac{\delta\left(\varepsilon(\mathbf{k}') - \varepsilon(\mathbf{k})\right)}{\left[2k^2(1-\cos\theta) + q_D^2\right]^2} \tag{2.80}$$

Where N_I is the number of impurities per unit volume; Z is number of charge units of the impurity; $\theta =$ angle between the initial and final wave-vectors $\mathbf{k}$ and $\mathbf{k}'$ respectively; ϵ_{si} is the permittivity of silicon; and $1/q_D =$ is the *Debye length* given by

$$q_D = \sqrt{\frac{q^2 n_0}{\epsilon_{\mathrm{si}} k_B T_L}} \tag{2.81}$$

where n_0 is the equilibrium electron density. By expanding $S(\mathbf{k},\mathbf{k}')$ in Legendre polynomials one can derive (p. 74 of [51]):

$$\begin{aligned} \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{impurity}} &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\frac{\partial f_l^m(\varepsilon)}{\partial t}\right]_{\text{impurity}} Y_l^m \\ &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{f_l^m(\varepsilon)}{\tau_{l,\,\text{impurity}}(\varepsilon)} Y_l^m \end{aligned} \tag{2.82}$$

Since impurity scattering is elastic, it turns out that $1/\tau_{l=0\,\text{impurity}}(\varepsilon) = 0$. Therefore for $l = 0$ and $l = 1$ we get:

$$\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impurity}} = 0 \tag{2.83}$$

$$\left[\frac{\partial f_{l=1}^m(\varepsilon)}{\partial t}\right]_{\text{impurity}} = -\frac{f_{l=1}^m(\varepsilon)}{\tau_{l=1,\,\text{impurity}}(\varepsilon)} \tag{2.84}$$

The scattering rate is

$$\frac{1}{\tau_{l=1,\,\text{impurity}}(\varepsilon)} = c_{\text{impu}} N_I \frac{\gamma'(\varepsilon)}{\sqrt{m^*}\gamma^{3/2}(\varepsilon)}\left(\log\left(\frac{4\gamma(\varepsilon)}{\gamma_{q_D}} + 1\right)\right) - \frac{4\gamma(\varepsilon)}{4\gamma(\varepsilon) + \gamma_{q_D}}\Bigg) \tag{2.85}$$

Where

$$\gamma_{q_D} = \frac{\hbar^2 q_D^2}{2m^*} \qquad \text{and} \qquad c_{\text{impu}} = \frac{\sqrt{2}Z^2 q^4}{64\pi^2 \epsilon_{\mathrm{si}}^2} \tag{2.86}$$

2.5.5 Impact Ionization Scattering

The scattering rates presented so far were derived by plugging the quantum-mechanics-derived transition rate in Eq. (2.65). But for impact ionization, we will use a heuristic approach. This approach is similar to, though not exactly same as, the one in [24, 43, 51]. The difference is in the treatment of the scattering-in term. We have used physical arguments to seek an expression for impact ionization such that condition of soon-to-be-presented Eq. (2.96) is satisfied. The derivation is as follows.

Due to impact ionization at any given point in k-space we loose electron due to scattering-out, and gain electrons due to scattering-in. The net generation rate of electron at $\mathbf{k}$ is simply the sum over all states while accounting for spin.

$$\begin{aligned}\text{Net Generation (e}^-\text{/time)} &= \sum_{\text{all } e^- \text{ states}} \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{impact}} \\ &= 2 \sum_{\text{all } \mathbf{k}} \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{impact}} \qquad (2.87)\\ &= 2\,\frac{\Omega}{(2\pi)^3}\int_{\text{all } \mathbf{k}} \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{impact}} d^3\mathbf{k}\end{aligned}$$

The net generation rate per unit volume is

$$\begin{aligned}G_{\text{impact}} &= \frac{\text{Net Generation (e}^-\text{/time)}}{\Omega} \\ &= \frac{2}{(2\pi)^3}\int_{\text{all } \mathbf{k}} \left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{impact}} d^3\mathbf{k} \\ &= \frac{2}{(2\pi)^3}\int_{\text{all } \mathbf{k}} \left[\frac{\partial f_0^0(k)}{\partial t}\right]_{\text{impact}} Y_0^0\, d^3\mathbf{k} \qquad (2.88)\\ &= \frac{2}{(2\pi)^3}\int_{\varepsilon} g(\varepsilon) \left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}} Y_0^0 \; d\varepsilon \\ &= \int_{\varepsilon} h(\varepsilon) \left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}} Y_0^0 \; d\varepsilon\end{aligned}$$

Where we have used the relationship of Eq. (2.41) $\int B(k)\, d^3\mathbf{k} = \int g(\varepsilon)\mathcal{B}(\varepsilon)\, d\varepsilon$ and the relation $h(\varepsilon) = g(\varepsilon)/4\pi^3$ from Eq. (2.47). For the sake of brevity we define $\tilde{h}(\varepsilon) \equiv h(\varepsilon)Y_0^0 = h(\varepsilon)/\sqrt{4\pi}$.

The task now is to derive the expression for scattering term $[\partial f_0^0/\partial t]_{\text{impact}}$. That integrand in the last line of Eq. (2.88) has physical significance:

$$\tilde{h}(\varepsilon) \left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}} \Delta\varepsilon = \text{Net increase in electrons due to impact ionization in an energy interval from } \varepsilon \text{ to } \varepsilon + \Delta\varepsilon \qquad (2.89)$$

The net increase is the difference between scattering-in and scattering-out:

$$\tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}} = \tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{in}} - \tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{out}} \tag{2.90}$$

The scattering-out has a relaxation-type form [24]

$$\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{out}} = \frac{f_0^0(\varepsilon)}{\tau_{\text{impact}}(\varepsilon)} \tag{2.91}$$

The scattering-in at ε requires a little more thought. Electrons scattering in from two sources: (1) electrons at some higher energy loose energy by initiating impact ionization and enter states of energy ε; and (2) electrons in valence band which are dislodged from the lattice and gain energy. Fig. 2.3 on the next page illustrates the impact-ionization process. The initiating electron has energy ε'. When it impact ionizes, it dislodges an electron from the top of the valence band, which is at energy $-\varepsilon_g$. ε_g is the band gap, which is 1.15eV for silicon. These two electrons, after the impact ionization event, are assumed to possess the same final energy (p. 79 Liang [51], and [43] and references therein). By energy-conservation the total initial and final energy can be equated. Total energy of the two electrons before impact-ionization is $\varepsilon' + (-\varepsilon_g)$. The final energy, which is same for both of them, is $\varepsilon + \varepsilon = 2\varepsilon$. Equating the two we get $2\varepsilon = \varepsilon' + (-\varepsilon_g)$. Hence $\varepsilon' = 2\varepsilon + \varepsilon_g$.

Let us consider electrons that are scattering in form the higher energy. Electron scattering into the energy range ε to $\varepsilon + \Delta\varepsilon$ are coming from a range $\varepsilon' = 2\varepsilon + \varepsilon_g$ to $\varepsilon' + \Delta\varepsilon' = 2(\varepsilon + \Delta\varepsilon) + \varepsilon_g$. From where we get $\Delta\varepsilon' = 2\Delta\varepsilon$. Electrons scattering in at ε to $\varepsilon + \Delta\varepsilon$ are simply the ones scattering out from ε' to $\varepsilon' + \Delta\varepsilon'$:

$$\begin{aligned} \tilde{h}(\varepsilon')\Delta\varepsilon'\left[\frac{\partial f_0^0(\varepsilon')}{\partial t}\right]_{\text{impact}}^{\text{out}} &= \frac{f_0^0(\varepsilon')}{\tau_{\text{impact}}(\varepsilon')}\tilde{h}(\varepsilon')\Delta\varepsilon' \\ &= \frac{f_0^0(2\varepsilon + \varepsilon_g)}{\tau_{\text{impact}}(2\varepsilon + \varepsilon_g)}\tilde{h}(2\varepsilon + \varepsilon_g)\ 2\Delta\varepsilon \end{aligned} \tag{2.92}$$

These electrons, along with the electrons dislodged from valence band, become the scattering-in electrons at energy ε. Multiplying the above equation by two, to account for the electrons from valence band, we get

$$\begin{aligned} \tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{in}}\Delta\varepsilon &= 2\frac{f_0^0(2\varepsilon + \varepsilon_g)}{\tau_{\text{impact}}(2\varepsilon + \varepsilon_g)}\tilde{h}(2\varepsilon + \varepsilon_g)\ 2\Delta\varepsilon \\ &= 4\frac{f_0^0(2\varepsilon + \varepsilon_g)}{\tau_{\text{impact}}(2\varepsilon + \varepsilon_g)}\tilde{h}(2\varepsilon + \varepsilon_g)\ \Delta\varepsilon \end{aligned} \tag{2.93}$$

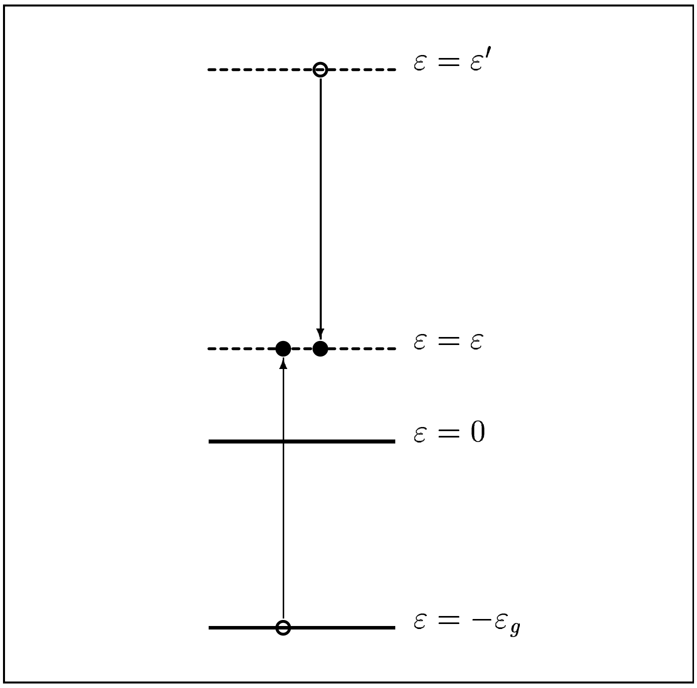

$\varepsilon = \varepsilon'$
$\varepsilon = \varepsilon$
$\varepsilon = 0$
$\varepsilon = -\varepsilon_g$

Figure 2.3: Illustration of the impact ionization process. An electron at energy ε' undergoes an impact-ionization event to dislodge an electron from the valence band at energy $-\varepsilon_g$. The two electrons, after impact-ionization, have the same energy ε. By energy-conservation, we have $\varepsilon' = 2\varepsilon + \varepsilon_g$.

Substituting Eq. (2.93) and Eq. (2.91) in Eq. (2.90) we get the scattering term we were seeking:

$$\boxed{g(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}} = 4\frac{f_0^0(2\varepsilon+\varepsilon_g)}{\tau_{\text{impact}}(2\varepsilon+\varepsilon_g)}g(2\varepsilon+\varepsilon_g) - g(\varepsilon)\,\frac{f_0^0(\varepsilon)}{\tau_{\text{impact}}(\varepsilon)}} \tag{2.94}$$

Where we used $g(\varepsilon)$ instead of $\tilde{h}(\varepsilon)$ since $\tilde{h}(\varepsilon) = (1/\sqrt{4\pi})(1/4\pi^3)g(\varepsilon)$. For The higher-order terms we can write

$$\boxed{\left[\frac{\partial f_l^m(\varepsilon)}{\partial t}\right]_{\text{impact}} = -\frac{f_l^m(\varepsilon)}{\tau_{\text{impact}}(\varepsilon)} \qquad \text{For } l \geq 1} \tag{2.95}$$

As a check we can see that we can calculate the number of particles generated per unit volume per unit time, G_{impact}, due to impact ionization by two means: integrating over the *total* rate or over the scattering-*out* part only. More specifically, it is easy to verify that the following equation is indeed true.

$$G_{\text{impact}} = \int_{\varepsilon} \tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{total}} d\varepsilon = \int_{\varepsilon} \tilde{h}(\varepsilon)\left[\frac{\partial f_0^0(\varepsilon)}{\partial t}\right]_{\text{impact}}^{\text{out}} d\varepsilon \tag{2.96}$$

We will, however, calculate the generation rate by the scattering-out part:

$$\begin{aligned} G_{\text{impact}} &= \int_{\varepsilon} \tilde{h}(\varepsilon) \left[\frac{\partial f_0^0(\varepsilon)}{\partial t} \right]_{\text{impact}}^{\text{out}} d\varepsilon \\ &= \int_{\varepsilon} \tilde{h}(\varepsilon) \frac{f_0^0(\varepsilon)}{\tau_{\text{impact}}(\varepsilon)} d\varepsilon \end{aligned} \tag{2.97}$$

The impact-ionization rate, $1/\tau_{\text{impact}}(\varepsilon)$ can be taken from any source, in this dissertation we use the rate given by Thoma *et al.* [60]:

$$\frac{1}{\tau_{\text{impact}}(\varepsilon)} = \begin{cases} 0 & \text{sec}^{-1} \text{ For } \frac{\varepsilon}{\text{eV}} < 1.128 \\ 0.875 \times 10^{13} \left(\frac{\varepsilon}{\text{eV}} - 1.128 \right)^3 & \text{sec}^{-1} \text{ For } 1.128 < \frac{\varepsilon}{\text{eV}} < 1.750 \\ 6.646 \times 10^{13} \left(\frac{\varepsilon}{\text{eV}} - 1.572 \right)^2 & \text{sec}^{-1} \text{ For } 1.750 < \frac{\varepsilon}{\text{eV}} \end{cases} \tag{2.98}$$

2.6 Final Form of the Boltzmann Equation

We collect all the scattering terms—after rewriting them in terms of $F_0^0(H)$ variable—together in the spherical-harmonic Boltzmann equation, Eq. (2.24):

$$\frac{\partial}{\partial x} \left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x} \right) + 3g \left[\frac{\partial F_0^0}{\partial t} \right]_{\text{collision}} = 0 \tag{2.99}$$

we get the final form

$$\begin{aligned} &\frac{\partial}{\partial x} \left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x} \right) && \text{(Space-Dependent)} \\ &\quad +3g \left\{ c_{\text{opt}} g^- \left(F_0^0(H - \hbar\omega_{\text{opt}}) - e^{\hbar\omega_{\text{opt}}/k_B T_L} F_0^0(H) \right) + \right. \\ &\quad\quad \left. c_{\text{opt}} g^+ \left(e^{\hbar\omega_{\text{opt}}/k_B T_L} F_0^0(H + \hbar\omega_{\text{opt}}) - F_0^0(H) \right) \right\} && \text{(Optical)} \\ &\quad +3g \left\{ c_{\text{ib}} g^{-(\nu')} \left(F_0^{0(\nu')}(H - \hbar\omega_{\text{ib}}) - e^{\hbar\omega_{\text{ib}}/k_B T_L} F_0^0(H) \right) + \right. \\ &\quad\quad \left. c_{\text{ib}} g^{+(\nu')} \left(e^{\hbar\omega_{\text{ib}}/k_B T_L} F_0^{0(\nu')}(H + \hbar\omega_{\text{ib}}) - F_0^0(H) \right) \right\} && \text{(Inter-Band)} \\ &\quad +3g c_{\text{ac}}^{\text{inelastic}} \sqrt{\gamma} \gamma'^2 \left\{ \left(1 + \frac{\gamma\gamma''}{\gamma'^2} \right) F_0^0 \right. \\ &\quad\quad \left. + \left(\frac{\gamma}{2\gamma'} + \left(1 + \frac{\gamma\gamma''}{\gamma'^2} \right) k_B T_L \right) \frac{\partial F_0^0}{\partial H} + \frac{\gamma k_B T_L}{2\gamma'} \frac{\partial^2 F_0^0}{\partial H^2} \right\} && \text{(Acoustic)} \\ &\quad +3 \left\{ 4g(2H + 2q\phi + \varepsilon_g) \frac{F_0^0(2H + q\phi + \varepsilon_g)}{\tau_{\text{impact}}(2H + 2q\phi + \varepsilon_g)} - \right. \end{aligned}$$

$$
\begin{aligned}
& g(H+q\phi)\ \frac{F_0^0(H)}{\tau_{\text{impact}}(H+q\phi)}\Bigg\} && \text{(Impact-Ionization)} \\
& = 0 && (2.100)
\end{aligned}
$$

The above equation is a self-adjoint *elliptic* differential equation. The form of the equation is same as that of the classical *diffusion* equations [113].

The total scattering rate is the sum of individual scattering rates:

$$
\frac{1}{\tau(\varepsilon)} = \frac{1}{\tau_{\text{opt}}(\varepsilon)} + \frac{1}{\tau_{\text{ib}}(\varepsilon)} + \frac{1}{\tau_{\text{ac}}(\varepsilon)} + \frac{1}{\tau_{l=1,\,\text{impurity}}(\varepsilon)} + \frac{1}{\tau_{\text{impact}}(\varepsilon)} \tag{2.101}
$$

2.7 Boundary Conditions

In this dissertation there are three type of boundaries that will be encountered for the spherical-harmonic-expanded Boltzmann equation: [1]

1. Ohmic boundary condition (Dirichlet type)
2. Insulator boundary condition (zero Neumann type)
3. Energy boundary condition (mixed type)

2.7.1 Ohmic Boundary Condition

Ohmic contact are present in all semiconductor devices. Non-rectifying metal contacts to the semiconductor fall in this category. They are means by which the currents and voltages are applied by the external world to the semiconductor device.

The physics of an ohmic contact is very complicated and not well understood, there are still controversies about the proper ohmic boundary condition to use in device simulation [63, 67–69]. In the absence of physical insight the exact mathematical statement of the boundary condition is not possible. We can, however, use engineering judgment to construct the boundary condition. For most semiconductor devices, ohmic contact are:

- Not part of the *active* or *intrinsic* region of the device,
- Sufficiently far from the active region,
- Near highly-doped electric-field-free low-average-velocity region of semiconductor.

---

[1] There are potentially more types of boundaries that we haven't considered in this dissertation, for example rectifying Schottky contacts.

The ohmic contact, therefore, is not expected to affect the intrinsic behavior of the semiconductor device. The ohmic contact in a MOSFET transistor, for instance, is far away from the the active region (the channel under the oxide) and is also not part of the intrinsic transistor action either. Since the semiconductor region close to the contact is usually highly-doped and has very low electric field. The average electron velocity is low, and we can therefore assume the contact to be in equilibrium.

The contact is believed to be an infinite supply or sink of electrons. We assume that at the ohmic contact electrons are at equilibrium. The energy distribution at equilibrium is Maxwellian: it has the form $\exp(-\varepsilon/k_B T_L)$. The boundary condition for the distribution is, therefore, of Dirichlet type:

$$f_0^0(\varepsilon) = C_0 \, e^{-\varepsilon/k_B T_L} \tag{2.102}$$

The constant C_0 is selected such that the total charge density, n, at the contact is the equilibrium density. Specifically, we set C_0 to satisfy $\int f_0^0(\varepsilon)\tilde{h}(\varepsilon)\, d\varepsilon = n$. The electron concentration n is evaluated by by imposing charge neutrality at the contact

$$p - n - N_D^+ - N_A^- = 0 \tag{2.103}$$

where p is the hole concentration, and N_D^+ and N_A^- are the ionized donor and acceptor doping concentrations respectively. Using the law of mass action, $np = n_i^2$, we get the electron charge density at equilibrium:

$$n = \frac{\mathcal{D} + \sqrt{\mathcal{D}^2 + 4n_i^2}}{2} \tag{2.104}$$

where $\mathcal{D} = N_D^+ - N_A^-$ is the net doping concentration.

A Dirichlet-type boundary condition sets the solution at the boundary independent of the solution inside the semiconductor device. This provides an element of stability to the numerical solution, and is therefore quite desirable from a numerical standpoint (page 191 [79]). Dirichlet conditions also provide the one equation at which the inequality of Eq. (4.60) holds trues, this is required for diagonal dominance.

2.7.2 Insulator Boundary Condition

Insulating boundaries in semiconductor devices occur in two situations: (a) in the presence of semiconductor-insulator interface, and (b) artificial boundaries in a device. A classic example of the first type is the semiconductor-oxide interface in a MOSFET. Oxide is an insulator and does not allow electrons to conduct. [2] Artificial boundaries are used in simulations to isolate the device from rest of the chip. Imposing the condition that no current flows from the device through that boundary ensures that the device being analyzed is electrically isolated, and hence can be simulated in isolation from rest of the chip.

An insulating boundary, therefore, is defined as an interface through which the electron current density is zero. From Eq. (2.63) the current density is $J = c_j \int \mathcal{J}_\varepsilon \, dH$ where $\mathcal{J}_\varepsilon = -g u_g^2 \tau \; \partial F_0^0/\partial x$. For zero current flow at all energies we have to impose the condition $\mathcal{J}_\varepsilon \equiv 0$. This gives the boundary condition

$$\frac{\partial F_0^0}{\partial x} = 0 \qquad (2.105)$$

In this boundary condition the derivative of the unknown is set to zero, therefore this is known as zero Neumann boundary condition. It is used at the silicon-oxide interface and at the artificial boundary of the MOSFET.

2.7.3 Energy Boundary Condition

The boundary condition in energy for the spherical-harmonic Boltzmann equation is tricky and interesting. The first thing to notice is that the spherical-harmonic Boltzmann equation, Eq. (2.24), does not have derivative in energy or Hamiltonian H. But the domain of the equation is curvilinear, so the derivative $\partial F_0^0/\partial x$ is still needed at the energy boundary. The energy boundary can be either the lower boundary, $\varepsilon = 0$ $(H = -q\phi(x))$, or the upper edge of the band, $\varepsilon = \varepsilon_{\max}$ $(H = \varepsilon_{\max} - q\phi(x))$. The boundary condition to be used is not immediately obvious. Lin *et al.* [32, 33, 48] in a similar situation, used a *substitute* boundary-condition derived by imposing current-continuity.

To derive the boundary condition we can use our knowledge of (a) mathematical structure of the equations, and (b) physical insight. The spherical-harmonic Boltzmann equation is

$$\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x}\right) + 3g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \qquad (2.106)$$

The term in the parentheses, $(\tau u_g^2 g \, \partial F_0^0/\partial x)$, is the differential electron current flux in x direction. If we draw analogy from heat transfer, $\partial F_0^0/\partial x$ is the driving force (similar to a temperature gradient), and $\tau u_g^2 g$ is a coefficient (similar to heat conductivity). This "conductivity" coefficient is a function of energy. As

---

[2] This, however, is strictly not true. If the electrons are very energetic, they can surmount the oxide barrier and be injected into the oxide. This, however, happens for only a very small fraction of electrons, therefore oxide can be safely considered to be an insulator

$\varepsilon \to 0$ the conductivity $\tau u_g^2 g \sim \varepsilon^{3/2} \to 0$, because $\tau \sim$ constant, $u_g \sim \sqrt{\varepsilon}$, and $g \sim \sqrt{\varepsilon}$. This means that there is no flux of electrons at $\varepsilon = 0$. This is satisfying from a physical and intuitive perspective too. There are no electron states at $\varepsilon < 0$, so an electron cannot leave the region $0 \leq \varepsilon \leq \varepsilon_{\max}$ to go to the $\varepsilon < 0$ region. So, referring to Fig. 2.4 on page 40, there will be no electron flux at point B.

This restriction on the flux at $\varepsilon = 0$ still does not directly translate into a boundary condition because we still do not know $\partial F_0^0/\partial x$, all we know is that the pre-multiplier, or conductivity, $\tau u_g^2 g$ goes to zero. To derive the boundary condition let us refer again to Fig. 2.4 on page 40. If we integrate Eq. (2.106) from A (at $x_0 - \Delta x$) to B (at x_0)

$$\int_A^B \frac{\partial}{\partial x}\left(\kappa \frac{\partial F_0^0}{\partial x}\right) dx + \int_A^B 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} dx = 0 \tag{2.107}$$

where we have set $\kappa = \tau u_g^2 g$. For a small Δx we can write

$$\left(\kappa \frac{\partial F_0^0}{\partial x}\right)_B - \left(\kappa \frac{\partial F_0^0}{\partial x}\right)_A + 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} \Delta x = 0 \tag{2.108}$$

Where $\kappa_B = \kappa(x_0) = 0$ because energy $\varepsilon = 0$. If we demand regularity and boundedness of the solution, the product $\kappa(\partial F_0^0/\partial x)$ is zero at point B. Eq. (2.108) then becomes

$$-\left(\kappa \frac{\partial F_0^0}{\partial x}\right)_A + 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} \Delta x = 0 \tag{2.109}$$

Which can be rewritten as

$$-\frac{\kappa_A}{\Delta x}\left(\frac{\partial F_0^0}{\partial x}\right)_A + 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{2.110}$$

$$\frac{0 - \kappa_A}{\Delta x}\left(\frac{\partial F_0^0}{\partial x}\right)_A + 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{2.111}$$

$$\frac{\kappa_B - \kappa_A}{\Delta x}\left(\frac{\partial F_0^0}{\partial x}\right)_A + 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{2.112}$$

In the limit $\Delta x \to 0$ $(\kappa_B - \kappa_A)/\Delta x = (\kappa(x_0) - \kappa(x_0 - \Delta x))/\Delta x = \partial\kappa/\partial x$. We then get the boundary condition in energy:

$$\boxed{\frac{1}{3g}\frac{\partial \kappa}{\partial x}\left(\frac{\partial F_0^0}{\partial x}\right) + \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0} \tag{2.113}$$

For the sake of physical insight, we derived the boundary condition equation in a slightly round-about way—first integrating then taking the limits to get a differential equation. We could have, as easily, just started from Eq. (2.106) and obtained

$$\frac{\partial}{\partial x}\left(\tau u_g^2 g\right)\left(\frac{\partial F_0^0}{\partial x}\right)+\tau u_g^2 g\left(\frac{\partial^2 F_0^0}{\partial x^2}\right)+3g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}}=0 \tag{2.114}$$

As $\varepsilon \to 0$ the coefficient of $\partial^2 F_0^0/\partial x^2$ goes to zero faster (as $\varepsilon^{3/2}$) than the other two terms (as $\varepsilon^{1/2}$). The second-order term, therefore, goes to zero faster than the others, and can be neglected. The other two terms tend to zero with the same power, so we can retain them; and, the resulting equation is the boundary condition. This equation is identical to Eq. (2.113).

This boundary condition involves the the value of the the variable F_0^0 (through the collision term) and its first derivative—it is, therefore, classified as a *mixed* boundary condition. A few comments about this boundary condition:

1. From the theory of differential equations, degenerate elliptic equations—equations whose highest-order coefficients tend to zero—do not need boundary conditions. It is easy to verify that the spherical-harmonic expanded Boltzmann equation, Eq. (2.106), is a degenerate elliptic equation.
2. The boundary condition is contained *within* the spherical-harmonic Boltzmann equation—we used the equation itself to derive a boundary condition for it. This, although initially a little odd, is intuitively satisfying for the following reason. The spherical-harmonic Boltzmann equation describes electron transport. At an energy boundary $\varepsilon = 0$ this equation simply says that the flux of electrons into the region $\varepsilon < 0$ is zero—we do not, and should not, need an external condition to tell us that; this information is, and should be, contained in our transport equation.
3. In [32, 33, 48] a substitute boundary condition was used at lower energy, which was a statement of conservation of electron current density. Boundary condition at higher energy, from physical arguments, was Dirichlet: $f_0^0(\varepsilon_{\max}) = 0$.
4. In our formulation too, we can set $f_0^0(\varepsilon_{\max}) = 0$. But this is not necessary; we can use the energy boundary condition Eq. (2.113) at $\varepsilon = \varepsilon_{\max}$ also.
5. The original Boltzmann equation, Eq. (2.5), does not have to worry about the $\varepsilon = 0$ boundary—this boundary does not exist since the independent variable, wave-vector $\mathbf{k}$, extends from $+\infty$ to $-\infty$. Spherical-harmonic expansion has introduced the magnitude of wave-vector $k = |\mathbf{k}|$ as the independent variable, thus creating a new boundary which did not exist before.

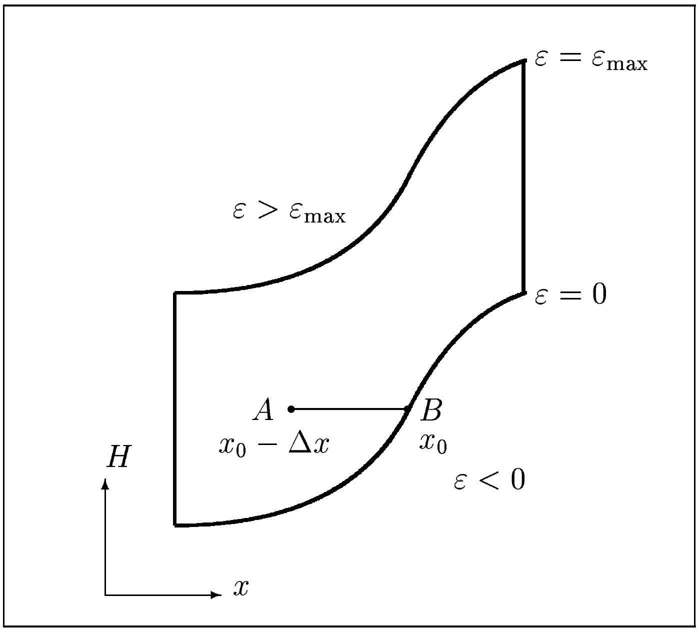

$\varepsilon = \varepsilon_{\max}$
$\varepsilon > \varepsilon_{\max}$
$\varepsilon = 0$
A
B
$x_0 - \Delta x$
x_0
H
$\varepsilon < 0$
x

Figure 2.4: Boundary condition in energy is derived by considering points A at $x_0 - \Delta x$, and B at x_0. In a given potential $\phi(x)$ the domain of the spherical-harmonic differential equation in x-H is $-q\phi(x) \leq H \leq \varepsilon_{\mathsf{max}} - q\phi(x)$ and $L_{\mathsf{min}} \leq x \leq L_{\mathsf{max}}$.

6. A boundary condition is normally a constraint the external world sets at a boundary. But the energy boundary is not available to us to apply any constraint. This, perhaps, is the physical reason why the energy boundary condition is contained within the Boltzmann equation itself.

7. Philosophically, we could go a step beyond. We can make a bold statement that there is no boundary in energy. The domain of the spherical-harmonic-expanded Boltzmann equation, as shown in Fig. 2.5 on the facing page, is in the entire region $L_{\min} \leq x \leq L_{\max}$; and, the governing equation is valid everywhere in the domain. For $\varepsilon \geq \varepsilon_{\max}$ or $\varepsilon \leq 0$, then, we have $\gamma(\varepsilon) = 0$ which sets variables like $g(\varepsilon)$ and $\kappa(\varepsilon) = \tau u_g^2 g$ to zero.

2.8 Chapter Summary

In this chapter the basic equations, that will be used throughout the dissertation, were introduced. The chapter first gave a brief derivation of the Boltzmann transport equation, and then outlined the spherical-harmonic-expansion approach to simplify it. Following that, some results pertaining to average quantities in the spherical-harmonic expansion of the distribution functions

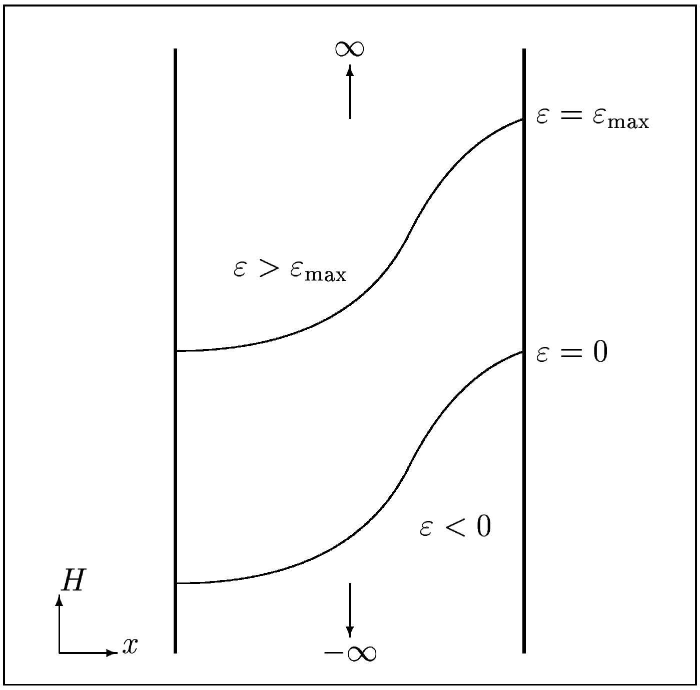

∞
$\varepsilon = \varepsilon_{\max}$
$\varepsilon > \varepsilon_{\max}$
$\varepsilon = 0$
$\varepsilon < 0$
H
x
$-\infty$

Figure 2.5: An alternate view of the energy boundary condition. We could consider the domain of the governing spherical-harmonic Boltzmann equation to be just $L_{\text{min}} \le x \le L_{\text{max}}$, with $\gamma(\varepsilon) = 0$ for $\varepsilon \ge \varepsilon_{\text{max}}$ or $\varepsilon \le 0$.

were derived. The scattering terms—which consist of optical, inter-band, acoustic, ionized impurity, impact ionization—were discussed. The boundary conditions for the resulting differential-difference equation were presented.

CHAPTER 3

MULTI-BAND BAND-STRUCTURE

3.1 Chapter Introduction

In the previous chapter basic expressions for the spherical-harmonic expansion of the Boltzmann equation were presented. We purposely swept the band-structure under the carpet, and discussed the derivation for a single band only. The two most important ingredients of device simulation by Boltzmann equation are (a) band-structure, and (b) scattering mechanisms. This chapter introduces the multi-band band-structure, and the scattering and transport constants.

In this dissertation a multi-band band-structure, with four conduction bands, by Brunetti *et al.* [57], and its improvement [58], is used. This band-structure has gained considerable popularity in spherical harmonic work [21, 23, 25, 31, 35–37, 39, 42, 43, 46, 50, 51]. Most spherical-harmonic simulations used only the lower two bands [21,31,35–37,39,42,43,46,51], and only a few incorporated all four bands [23, 25, 50]—in fact, all four bands have been utilized only in space-independent simulations. In contrast, this dissertation incorporates all four bands in space-dependent formulations: both in one-dimensional and two-dimensional formulations. This is a salient feature of the dissertation.

The chapter is organized as follows. The multi-band band-structure is introduced in Section 3.2 on the next page. The Boltzmann equation in this band-structure is discussed in Section 3.3 on page 46. It is shown there that we need to solve only two Boltzmann equations in this band-structure. This band-structure also introduces some changes to the scattering term, these are discussed in Section 3.4 on page 48.

3.2 Multi-Band Band-Structure

Brunetti *et al.*, 1989, proposed a new multi-band band-structure for silicon. If ε is the energy of electrons, and k the magnitude of the electron wave-vector, then the relationship between ε and k is represented by the relations:

$$\varepsilon + \alpha\varepsilon^2 = \frac{\hbar^2 k^2}{2m^{*(1)}} \qquad \text{(Band 1)} \qquad (3.1)$$

$$\varepsilon = \varepsilon_{\max}^{(2)} - \frac{\hbar^2 k^2}{2m^{*(2)}} \qquad \text{(Band 2)} \qquad (3.2)$$

$$\varepsilon = \varepsilon_{\min}^{(3)} + \frac{\hbar^2 k^2}{2m^{*(3)}} \qquad \text{(Band 3)} \qquad (3.3)$$

$$\varepsilon = \varepsilon_{\max}^{(4)} - \frac{\hbar^2 k^2}{2m^{*(4)}} \qquad \text{(Band 4)} \qquad (3.4)$$

where $\hbar$ is the usual reduced Planck's constant; α is the non-parabolicity factor; and m^* is the effective mass; and constants $\varepsilon_{\max}^{(2)}$, $\varepsilon_{\min}^{(3)}$, $\varepsilon_{\max}^{(4)}$ are the energy limits of the respective bands. To make the spherical-harmonic derivations more general we express the above relations with a *dispersion relation* for each band:

$$\boxed{\gamma^{(\nu)}(\varepsilon) = \frac{\hbar^2 k^2}{2m^{*(\nu)}}} \qquad \text{For } \nu = 1,\ 2,\ 3, \text{ and } 4 \qquad (3.5)$$

The specific form for each band is

$$\gamma^{(1)}(\varepsilon) = \varepsilon + \alpha\varepsilon^2 \qquad \text{(Band 1)} \qquad (3.6)$$

$$\gamma^{(2)}(\varepsilon) = \varepsilon_{\max}^{(2)} - \varepsilon \qquad \text{(Band 2)} \qquad (3.7)$$

$$\gamma^{(3)}(\varepsilon) = \varepsilon - \varepsilon_{\min}^{(3)} \qquad \text{(Band 3)} \qquad (3.8)$$

$$\gamma^{(4)}(\varepsilon) = \varepsilon_{\max}^{(4)} - \varepsilon \qquad \text{(Band 4)} \qquad (3.9)$$

The ε-k relation of this band-structure is shown in Fig. 3.1 on the facing page, and the numerical values are given in Table 3.1 on page 46.

This band-structure very closely approximates the density of states (DOS) obtained from band-structure calculations [57]. In addition this band-structure mimics the details of the actual band-structure of silicon: first two bands (1 & 2) reproduce the details along < 100 > crystallographic direction; and upper two bands (3 & 4) reproduce details in the < 110 > direction [57,60,61].

Using the expression for density of states, $h(\varepsilon)$, from Eq. (2.47) we can calculate the total density of states for our band-structure

$$DOS = Z^{(1)}h^{(1)}(\varepsilon) + Z^{(2)}h^{(2)}(\varepsilon) + Z^{(3)}h^{(3)}(\varepsilon) + Z^{(4)}h^{(4)}(\varepsilon) \qquad (3.10)$$

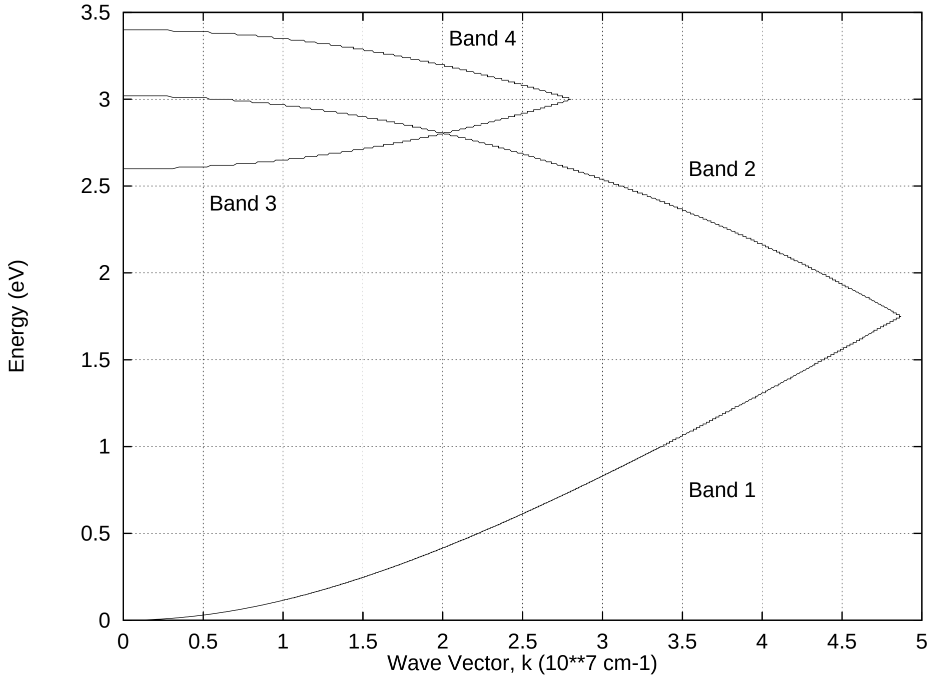


Figure 3.1: The multi-band band-structure of Brunetti *et al.* [57] used in this dissertation. This spherical band-structure has four bands: lowest band is non-parabolic, and the rest are parabolic. Bands 1 and 3 are electron-like (increasing density of states) and bands 2 and 4 are hole-like (decreasing density of states).

Band	m^*/m_0	ε_{min} (eV)	ε_{max} (eV)	Z	α eV^{-1}
1	0.320	0	1.75	6	0.35
2	0.712	1.75	3.02	6	
3	0.750	2.60	3.00	12	
4	0.750	3.00	3.40	12	

Table 3.1: Numerical values for the four bands of the band-structure [57]. Effective mass m^* is given in terms of the electron mass m_0. Each band extends from ε_{min} to ε_{max}. Z is the band multiplicity, or the number of equivalent symmetrical bands. α is the non-parabolicity factor.

where $Z^{(\nu)}$ is the band multiplicity, or the number of equivalent symmetrical bands, in band (ν). In Fig. 3.2 on the facing page we notice that density of states increases in the energy range 0 to 1.75 and 2.6 to 3eV; and it decreases from 1.75 to 2.6 and 3 to 3.4eV.

3.3 Boltzmann Equation in Multi-Band Band-Structure

When a band-structure has multiple bands we can think of electrons populating each band as a distinct species—each band is described by a unique Boltzmann equation, which is coupled to other Boltzmann equations by inter-band scattering. We must therefore write a Boltzmann equation for all the four bands, as well all their equivalent bands. It is, of course, always desirable to have a parsimonious set of Boltzmann equations.

Since our multi-band band-structure has four bands with multiplicity, we look for ways to reduce the number of Boltzmann equations. The first saving comes from noting that electrons in equivalent valleys of the same band behave identically, therefore one Boltzmann equation per band is sufficient.

The second saving comes by treating the bands (1) and (2) as one band, which permits us to write only one Boltzmann equation for both. This has been justified in reference [21]. At the boundary of bands (1) and (2) energy and group velocity is continuous. Electrons can go from band (1) to (2) (and vice versa) due to acceleration in electric field, in addition to inter-band scattering [60]. One unique distribution function, therefore, describes population of both bands. Same holds true for bands (3) and (4). Notationally we refer to the the concatenated bands (1) and (2) by the symbol (12), and bands (3) and (4) by the symbol (34).

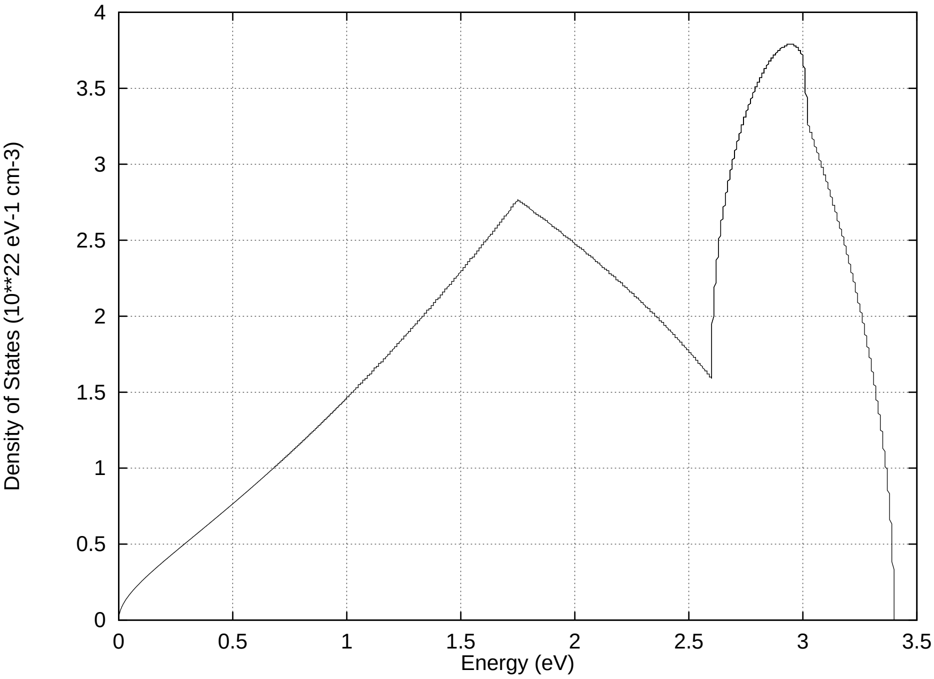


Figure 3.2: Density of states for the present multi-band band-structure. The parameters of this simplified band-structure are chosen to closely approximate the density of states of silicon by other calculations.

The spherical-harmonic Boltzmann equation of Eq. (2.24) for band (12) is

$$\frac{1}{3g^{(12)}}\frac{\partial}{\partial x}\left(\kappa^{(12)}\frac{\partial F_0^{0\,(12)}}{\partial x}\right)+\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{coll}}=0 \tag{3.11}$$

where we used the symbol $\kappa^{(12)} = \tau^{(12)}\left(u_g^{(12)}\right)^2 g^{(12)}$ and the collision term is

$$\begin{aligned}\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{coll}} = &\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{opt}}+\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{ib}}^{(12)\leftrightarrow(34)}+\\ &\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{ac}}+\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{impurity}}+\left[\frac{\partial F_0^{0\,(12)}}{\partial t}\right]_{\text{impact}}\end{aligned} \tag{3.12}$$

Since ionized impurity scattering is elastic, it need not be mentioned in the above formula. A similar Boltzmann equation can be written for band (34): we would interchange the superscripts (12) and (34) in Eq. (3.11) and Eq. (3.12). The inter-band collision term couples the two Boltzmann equations.

3.3.1 Total Distribution

Expressing the distribution for the concatenated bands (1) and (2) is simple since the two bands are non-overlapping. We can also define a *total* distribution for all the bands as a weighted sum of density of states:

$$f_0^{0\,(\mathrm{Total})} = \frac{Z^{(12)} h^{(12)}(\varepsilon) f_0^{0\,(12)} + Z^{(34)} h^{(34)}(\varepsilon) f_0^{0\,(34)}}{Z^{(12)} h^{(12)}(\varepsilon) + Z^{(34)} h^{(34)}} \tag{3.13}$$

Where $Z^{(12)} = Z^{(1)} = Z^{(2)}$ and $Z^{(34)} = Z^{(3)} = Z^{(4)}$.

3.4 Collision Term and Transport Model

The band-structure of Section 3.2 on page 44 needs transport parameters to completely define electron transport. These transport parameters are not unique, but are determined by numerical experiments to match experimental results. Depending on which device we intend to simulate, we may tune our parameters accordingly. The original paper by Brunetti *et al.* [57] contained transport parameters suitable for bulk Silicon and low-field transport. Fiegna and Sangiorgi [58] adopted the same band-structure, but used a larger set of experiments to determine the transport parameters. They proposed two set of parameters, one geared towards MOS (metal-oxide-semiconductor) devices and other more suitable for bipolar devices. The transport parameters are listed in Table 3.2 on the facing page. In the dissertation we use the MOS parameters for simulating the $\mathrm{n}^+\mathrm{nn}^+$ device and MOSFET's, whereas the bipolar parameters were used for the bipolar transistor. The $\mathrm{n}^+\mathrm{nn}^+$ device mimics a MOS device, therefore, it inherits the MOS parameters.

The treatment of scattering in Section 2.5 on page 26 was a general treatment for any band. The multi-band band-structure necessitates us to re-visited and embellish that treatment.

3.4.1 Optical Phonon Scattering

In Section 2.5.1 on page 27 the expressions for optical phonon were derived without considering band multiplicity. Therefor we multiply the transition rate of Eq. (2.67) by Z, the multiplicity of the band. $Z = 6$ for bands (1) and (2), and $Z = 12$ for bands (3) and (4). The factor Z would then be also find its way to Eq. (2.71). Section II of reference [58] explains this in more detail. Fig. 3.3 on page 50 plots the optical phonon scattering rate.

	MOS	
$\mathcal{D}_{\text{opt}}$	5×10^8 eV/cm	Optical deformation potential
$\hbar\omega_{\text{opt}}$	0.0517eV	Optical phonon energy ($T_{\text{opt}} = 600$K)
		(Same for inter-band optical phonon)
$\mathcal{D}_{\text{ac}}$	8.3eV ($\varepsilon < 0.15$eV)	Acoustic deformation potential
	4.0eV ($\varepsilon > 0.35$eV)	(Linear for 0.15eV $< \varepsilon <$ 0.35eV)
v_{sound}	9×10^5 cm/s	Sound velocity
ξ_{impurity}	2.5	Impurity scattering enhancement factor
P_{impact}	4.0×10^{-3}	Impact ionization pre-factor

	Bipolar	
$\mathcal{D}_{\text{opt}}$	4.9×10^8 eV/cm	Optical deformation potential
P_{impact}	5.0×10^{-2}	Impact ionization pre-factor

Table 3.2: The values of transport parameters from Fiegna and Sangiorgi [58]. The first set of parameters is geared towards MOS (metal-oxide-semiconductor) devices, and the second set is more suitable for bipolar devices. For the bipolar set only those parameters that are different from the MOS set are mentioned.

3.4.2 Inter-Band Optical Phonon Scattering

The inter-band optical phonons scattering too is modified in a manner similar to the optical phonons scattering. The c_{ib} in Eq. (2.75) is multiplied by the multiplicity of final valleys Z_f. For inter-band scattering between bands (1) and (2) $Z_f = Z-1 = 6-1$; and between bands (3) and (4) $Z_f = Z-1 = 12-1$. Scattering from band (12) to (34) we will have $Z_f = Z = 12$; and scattering from band (34) to (12) will have $Z_f = Z = 6$. This too has been explained in section II of reference [58].

3.4.3 Acoustic Phonon Scattering

Acoustic scattering formulas remain the same, but the only salient difference is that the acoustic phonon deformation potential, $\mathcal{D}_{\text{ac}}$, depends on energy, as indicated in Table 3.2. Therefore the acoustic scattering rate, as shown in Fig. 3.3 on the next page has a “hump” at low energies. It may also be seen that, except at low energy, acoustic phonon scattering rate is much smaller than optical phonon rate.

3.4.4 Ionized Impurity Scattering

To obtain better agreement with experimental data, Fiegna and Sangiorgi [58] multiply the Debye screening length, Eq. (2.81), by a factor, ξ_{impurity}, called the

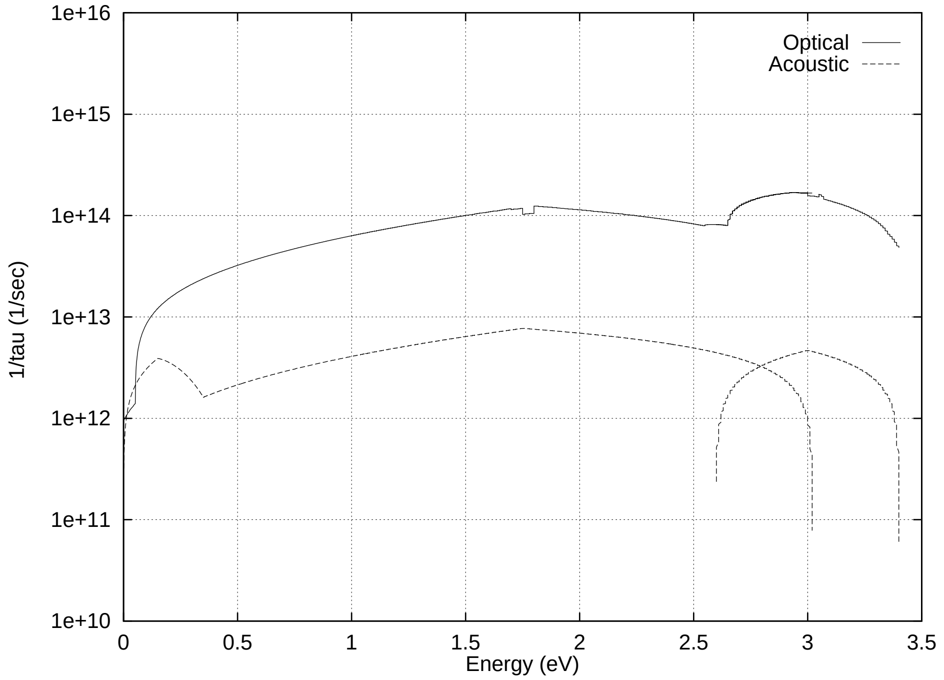


Figure 3.3: Optical and acoustic phonon scattering rate.

impurity scattering enhancement factor. The scattering rate for two impurity doping concentrations is plotted in Fig. 3.4 on the facing page.

3.4.5 Impact Ionization Scattering

To match experimental data that depends on impact ionization (ionization coefficient, substrate current) the impact ionization scattering rate by Thoma *et al.* [60], Eq. (2.98) in Section 2.5.5, is multiplied by a pre-factor P_{impact} [58].

$$\frac{1}{\tau_{\text{impact}}(\varepsilon)} = P_{\text{impact}} \frac{1}{\tau_{\text{impact}}^{\text{thoma}}(\varepsilon)} \tag{3.14}$$

This scattering rate, with $P_{\text{impact}} = 1$, is plotted in Fig. 3.5 on page 52.

In addition we notice that for every lower band there are two upper bands ($Z = 6$ for lower bands, while it is $Z = 12$ for upper bands). Therefore, the scattering-in from the upper bands is multiplied by two.

3.5 Chapter Summary

In this chapter the multi-band band-structure was introduced. Boltzmann equations were written in this band-structure and it was indicated that the

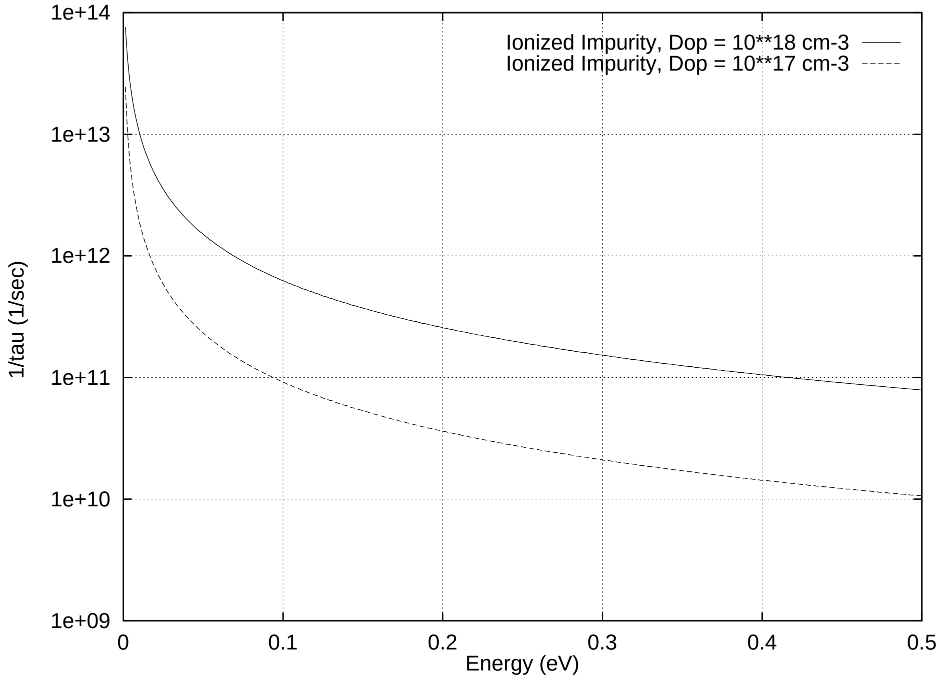


Figure 3.4: Ionized impurity scattering rate for two impurity doping concentrations: $N_I = 10^{18}$ and 10^{17} cm^{-3}. The equilibrium electron concentration is assumed to be equal to the impurity concentration: $n_0 = N_I$.

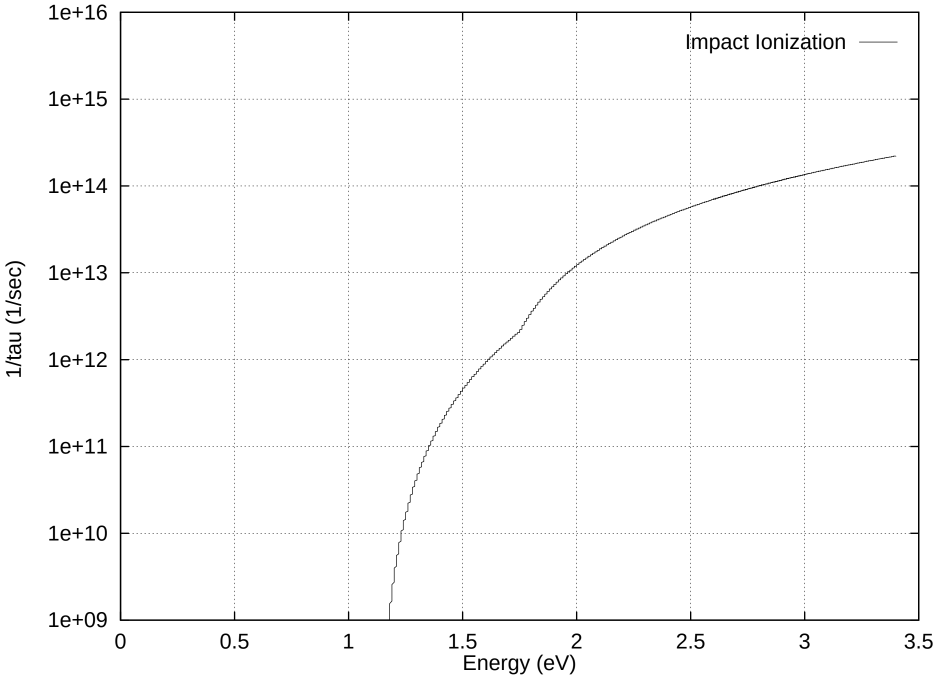


Figure 3.5: Impact ionization scattering rate of Thoma *et al.* [60] with unity pre-factor.

number of Boltzmann equations can be reduced to two. Due to considerations arising from multi-band band-structure, some changes necessary in scattering terms were discussed. The transport parameters suitable for MOS (metal-oxide-semiconductor) and bipolar devices were introduced.

CHAPTER 4

TRANSFORMATION AND DISCRETIZATION OF THE BOLTZMANN TRANSPORT EQUATION

4.1 Chapter Introduction

In Chapter 2, the original Boltzmann equation (Eq. (2.5)) in wave-vectors was projected on spherical-harmonics to derive a simple, easy-to-solve Boltzmann equation involving the coefficients of the expansion (Eq. (2.24), or Eq. (4.3) in this chapter). This simplified equation is written in terms of the energy distribution function [31, 35–37, 39, 41–43, 46, 49–51]. A close examination of this equation reveals that it leads to discretized equations that do not have the desirable property of diagonal dominance. Spoilers of diagonal dominance are the scattering terms, especially optical phonons. In addition, diagonal dominance is worse in the important low-energy region of the distribution—average quantities are more sensitive to the distribution at low energies.

In this dissertation a new variable-transformation of the Boltzmann equation is proposed: Boltzmann equation is transformed from the original variable of energy distribution to a new pre-Maxwellian variable.This is a salient contribution of this dissertation. Devices are simulated by discretization of the transformed Boltzmann equation. The new pre-Maxwellian variable has several desirable properties. (a) It enhances the Diagonal dominance of the discretized equations; the optical, inelastic-acoustic, and inter-band optical phonons are transformed to diagonally-dominant forms. (b) It accounts for the rapid exponential variation of the distribution in both energy and space. (c) The new pre-Maxwellian variable allow us to write the electron concentrations in Slotboom variables, which opens the possibility of using superior Poisson-equation solvers like Mayergoyz' fixed-point algorithm [95–97].(d) These advantages accrue with no penalty in linearity: Boltzmann equation in the new variable is still linear. In fact, the structure of the spherical-harmonic Boltzmann equation before and after transformation is similar; therefore, old solvers, with minor modifications, can be used for the new variable.

Change of variables in device simulation is not a new idea; the drift-diffusion model has been cast in many variables: the potential and carrier-concentration variable; Slotboom variables; and quasi-Fermi potential variables [98,102,104,105]. Variable transformations have also been applied to the hydrodynamic device simulation model [51,55,56] and process simulation [112]. In fact, in spherical-harmonic approach we have transformed the equations from the energy variable to the Hamiltonian variable, as shown in Section 2.3 on page 14. Mathematical properties and numerical behavior is different for different variables [102], which can be exploited in designing the device simulator.

In addition, this chapter presents a new formulation for the inelastic acoustic phonons. This formulation is suitable for numerical discretization: it is self-adjoint and avoids first derivatives.

This chapter is organized as follows: Section 4.2 introduces the new pre-Maxwellian variable and discusses the transformation of the spherical-harmonic Boltzmann equation to the new variable. The discretization of the resulting equations is formulated in Section 4.3 on page 61, where the control-volume principles are enunciated and applied to the Boltzmann equation. Since the advantages of the new variable cannot be discussed until the discretization has been formulated, we wait till Section 4.4 on page 71 to discuss these advantages. Finally in Section 4.5 the chapter is summarized.

4.2 Transformation

4.2.1 The Pre-Maxwellian Variable

The spherical-harmonic-expanded Boltzmann equation, Eq. (2.24), was written in terms of the energy distribution function $F_0^0(x, H)$; we transform it to a new *pre-Maxwellian* variable:

$$\boxed{F_0^0(x, H) = C(x, H)\ e^{-H/k_B T_L}} \tag{4.1}$$

Since the Hamiltonian $H = \varepsilon - q\phi(x)$, the exponent term contains the $\exp(-\varepsilon/k_B T_L)$ Maxwellian factor. In addition, there is also a dependence on space-dependent electric potential, $\exp(q\phi(x)/k_B T_L)$. This variable transformation contains both the dependencies:

$$F_0^0(x,H) = C(x,H)\; e^{-H/k_BT_L} = C(x,H) \quad \underbrace{e^{-\varepsilon/k_BT_L}}_{\text{Energy Dependence}} \quad \underbrace{e^{q\phi(x)/k_BT_L}}_{\text{Space Dependence}} \tag{4.2}$$

Section 4.4 on page 71 discusses the motivation and advantages of this new variable. Next task, for now, is to apply this transformation to the Boltzmann equation and the scattering terms and derive a new equation in the unknown C.

4.2.2 Boltzmann Equation in Pre-Maxwellian Variable

The starting point is the spherical-harmonic Boltzmann equation, Eq. (2.24):

$$\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x}\right) + 3g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{4.3}$$

We substitute for the distribution function $F_0^0(x,H) = C(x,H)\cdot\exp(-H/k_BT_L)$ in every term of the above equation. Space-dependent term is easy to evaluate: the exponent factors out because the partial derivative is taken with H as constant.

$$\frac{\partial F_0^0}{\partial x} = \left.\frac{\partial F_0^0(x,H)}{\partial x}\right|_{\text{constant H}} = e^{-H/k_BT_L}\frac{\partial C(x,H)}{\partial x} \tag{4.4}$$

which leads to

$$\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x}\right) = \frac{\partial}{\partial x}\left(\tau u_g^2 g\; e^{-H/k_BT_L}\frac{\partial C}{\partial x}\right) = e^{-H/k_BT_L}\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial C}{\partial x}\right) \tag{4.5}$$

The space-dependent term, therefore, retains the original linear self-adjoint form. We redefine the collision term as

$$\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} \equiv e^{-H/k_BT_L}\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} \tag{4.6}$$

by introducing a new symbol $[\partial C/\partial t]_{\text{collision}}$. Substituting the last two equations, Eq. (4.6) and Eq. (4.5), into Eq. (4.3) we get

$$e^{-H/k_BT_L}\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial C}{\partial x}\right) + e^{-H/k_BT_L}\,3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} = 0 \tag{4.7}$$

which leads to the spherical-harmonic Boltzmann equation in the new pre-Maxwellian variable:

$$\boxed{\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial C}{\partial x}\right) + 3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} = 0} \tag{4.8}$$

This new equation is noticeably similar in form and linearity to the original equation, Eq. (4.3). We now turn our attention to the collision term to complete the derivation.

4.2.3 Collision Terms in Pre-Maxwellian Variable

We apply the variable transformation on the five scattering terms:

1. Optical Phonon Scattering
2. Inter-band Optical Phonon Scattering
3. Inelastic Acoustic Phonon Scattering
4. Ionized Impurity Scattering
5. Impact Ionization Scattering

4.2.3.1 Optical Phonon Scattering

We start from from the expression for optical phonon scattering of the Y_0^0 coefficient of the spherical-harmonic expansion (Eq. (2.68) in Section 2.5.1).

$$\begin{aligned}\left[\frac{\partial F_0^0(\varepsilon)}{\partial t}\right]_{\text{opt}} &= c_{\text{opt}} g^- \left(F_0^0(H-\hbar\omega_{\text{opt}}) - e^{\hbar\omega_{\text{opt}}/k_BT_L} F_0^0(H)\right) + \\ &\quad c_{\text{opt}} g^+ \left(e^{\hbar\omega_{\text{opt}}/k_BT_L} F_0^0(H+\hbar\omega_{\text{opt}}) - F_0^0(H)\right)\end{aligned} \tag{4.9}$$

Using the transformation $F_0^0(H) = C(H)\exp(-H/k_BT_L)$ in $F_0^0(H \pm \hbar\omega_{\text{opt}})$ we get

$$F_0^0(H+\hbar\omega_{\text{opt}}) = C(H+\hbar\omega_{\text{opt}}) e^{-(H+\hbar\omega_{\text{opt}})/k_BT_L} \tag{4.10}$$

$$F_0^0(H-\hbar\omega_{\text{opt}}) = C(H-\hbar\omega_{\text{opt}}) e^{-(H-\hbar\omega_{\text{opt}})/k_BT_L} \tag{4.11}$$

Above two expressions can be substituted in Eq. (4.9) to yield

$$\begin{aligned}\left[\frac{\partial F_0^0(\varepsilon)}{\partial t}\right]_{\text{opt}} &= c_{\text{opt}} g^- e^{-H/k_BT_L} e^{\hbar\omega_{\text{opt}}/k_BT_L} \left(C(H-\hbar\omega_{\text{opt}}) - C(H)\right) + \\ &\quad c_{\text{opt}} g^+ e^{-H/k_BT_L} \left(C(H+\hbar\omega_{\text{opt}}) - C(H)\right) \\ &= e^{-H/k_BT_L} \left[\frac{\partial C}{\partial t}\right]_{\text{opt}}\end{aligned} \tag{4.12}$$

This gives the expression for optical phonon scattering in the C variable

$$\begin{aligned}\left[\frac{\partial C}{\partial t}\right]_{\text{opt}} &= c_{\text{opt}} g^- e^{\hbar\omega_{\text{opt}}/k_BT_L} \left(C(H-\hbar\omega_{\text{opt}}) - C(H)\right) + \\ &\quad c_{\text{opt}} g^+ \left(C(H+\hbar\omega_{\text{opt}}) - C(H)\right)\end{aligned} \tag{4.13}$$

4.2.3.2 Inter-band Optical Phonon Scattering

Transforming the inter-band phonon scattering is very similar to transforming the optical phonons. We start from Eq. (2.72) in Section 2.5.2 and use the transformation $F_0^0(H) = C(H)\exp(-H/k_BT_L)$ to derive

$$\begin{aligned}\left[\frac{\partial C}{\partial t}\right]_{\text{ib}}^{(\nu\leftrightarrow\nu')} &= c_{\text{ib}}g^{-(\nu')}e^{\hbar\omega_{\text{ib}}/k_BT_L}\left(C^{(\nu')}(H-\hbar\omega_{\text{ib}}) - C^{(\nu)}(H)\right) + \\ &\quad c_{\text{ib}}g^{+(\nu')}\left(C^{(\nu')}(H+\hbar\omega_{\text{ib}}) - C^{(\nu)}(H)\right)\end{aligned} \tag{4.14}$$

4.2.3.3 Inelastic Acoustic Phonon Scattering

Readers familiar with the derivation and discretization of inelastic inelastic acoustic phonons in previous spherical-harmonic work [20, 32, 33, 47, 48] may notice the difference in the derivation here. In this sub-section a new derivation for the inelastic acoustic phonons is presented which results in a novel self-adjoint formulation. The guiding aim is to seek self-adjoint second-order derivatives; and, if possible, shun first-derivatives [113]. [1] Let us start form the expression for inelastic acoustic phonons, Eq. (2.77) in Section 2.5.3. If we rewrite the expression using $F_0^0(H)$ instead $f_0^0(\varepsilon)$ we get

$$\begin{aligned}\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} &= c_{\text{ac}}^{\text{inelastic}}\sqrt{\gamma}\gamma'^2\left\{\left(1+\frac{\gamma\gamma''}{\gamma'^2}\right)F_0^0\right. \\ &\quad \left. + \left(\frac{\gamma}{2\gamma'} + \left(1+\frac{\gamma\gamma''}{\gamma'^2}\right)k_BT_L\right)\frac{\partial F_0^0}{\partial H} + \frac{\gamma k_BT_L}{2\gamma'}\frac{\partial^2 F_0^0}{\partial H^2}\right\}\end{aligned} \tag{4.15}$$

This formula is not very convenient to use in numerical discretization because it has first derivatives. First-derivative terms are rather difficult to handle and often are the cause of numerical problems. Second-derivative terms, on the other, hand lead to stable discretizations. With this in mind, one may notice that the collision term in the Boltzmann equation (Eq. (4.3) or Eq. (4.8)) never occurs by itself; it always occurs as a product of g and the collision integral, $g[\partial F/\partial t]_{\text{collision}}$. So, one might as well seek a derivation for the product, instead of just the collision term. In doing so, as we will soon see, we get a very convenient expression for the inelastic acoustic phonon scattering. There is a physical reason for that too. Since g is proportional to the density of states, the product $g[\partial F/\partial t]_{\text{collision}}\Delta H$ is proportional to the net increase in number of electrons in interval ΔH around H due to inelastic acoustic phonon scattering. Usually when we deal with *actual* electron populations the mathematical expression come out in self-adjoint form.

---

[1] Second-derivatives terms, after discretization, couple with the neighboring grid points in a manner suitable for numerical solution; they usually produce diagonally dominant matrices. First derivatives, on the other hand, have to be discretized very carefully; they often have trouble giving diagonally-dominant matrices, and are often plagued by numerical noise. A good example is the electron continuity equation in the drift-diffusion model, where the second-order self-adjoint term by itself poses no problem but the presence of the first-derivative term complicates the discretization: We are forced to use schemes like the exponential Scharfetter-Gummel [100], or upwind, or hybrid schemes [113, 114].

Multiplying Eq. (4.15) by $g = c_g\sqrt{\gamma}\gamma'$ and taking γ'^2 inside the brackets we get

$$\begin{aligned} g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} &= c_g c_{\text{ac}}^{\text{inelastic}}\gamma\gamma'\left\{\left(\gamma'^2+\gamma\gamma''\right)F_0^0\right. \\ &\quad \left. +\left(\frac{\gamma\gamma'}{2}+\left(\gamma'^2+\gamma\gamma''\right)k_BT_L\right)\frac{\partial F_0^0}{\partial H}+\frac{\gamma\gamma'}{2}k_BT_L\frac{\partial^2 F_0^0}{\partial H^2}\right\} \end{aligned} \tag{4.16}$$

By re-arranging the terms in the curly brackets

$$\begin{aligned} g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} &= c_g c_{\text{ac}}^{\text{inelastic}}\gamma\gamma'\left\{\left(\gamma'^2+\gamma\gamma''\right)\left(F_0^0+k_BT_L\frac{\partial F_0^0}{\partial H}\right)\right. \\ &\quad \left. +\frac{\gamma\gamma'}{2}\frac{\partial}{\partial H}\left(F_0^0+k_BT_L\frac{\partial F_0^0}{\partial H}\right)\right\} \end{aligned} \tag{4.17}$$

and using a temporary variable T for brevity,

$$T = F_0^0 + k_BT_L\frac{\partial F_0^0}{\partial H} = e^{-H/k_BT_L}\frac{\partial}{\partial H}\left(k_BT_Le^{H/k_BT_L}F_0^0\right) \tag{4.18}$$

we can rewrite Eq. (4.17) by noticing that $(\gamma\gamma')' = \gamma'^2 + \gamma\gamma''$.

$$\begin{aligned} g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} &= c_g c_{\text{ac}}^{\text{inelastic}}\left\{\gamma\gamma'\frac{\partial(\gamma\gamma')}{\partial H}T+\frac{(\gamma\gamma')^2}{2}\frac{\partial T}{\partial H}\right\} \\ &= c_g c_{\text{ac}}^{\text{inelastic}}\frac{\partial}{\partial H}\left(\frac{(\gamma\gamma')^2}{2}T\right) \\ &= \frac{c_g c_{\text{ac}}^{\text{inelastic}}}{2}\frac{\partial}{\partial H}\left((\gamma\gamma')^2\,e^{-H/k_BT_L}\frac{\partial}{\partial H}\left(k_BT_Le^{H/k_BT_L}F_0^0\right)\right) \\ &= \frac{c_g c_{\text{ac}}^{\text{inelastic}}}{2}k_BT_L\frac{\partial}{\partial H}\left((\gamma\gamma')^2\,e^{-H/k_BT_L}\frac{\partial}{\partial H}\left(e^{H/k_BT_L}F_0^0\right)\right) \\ &= c_{\text{ac, gC}}^{\text{inelastic}}\frac{\partial}{\partial H}\left((\gamma\gamma')^2\,e^{-H/k_BT_L}\frac{\partial C}{\partial H}\right) \end{aligned} \tag{4.19}$$

We introduced new constant $c_{\text{ac, gC}}^{\text{inelastic}} \equiv c_g c_{\text{ac}}^{\text{inelastic}}k_BT_L/2$. From our definition of collision term in C, Eq. (4.6), the expression for inelastic collision term in the C variable is derived:

$$\begin{aligned} g\left[\frac{\partial C}{\partial t}\right]_{\text{ac}} &= e^{H/k_BT_L} g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} \\ &= c_{\text{ac, gC}}^{\text{inelastic}}\, e^{H/k_BT_L} \frac{\partial}{\partial H}\left((\gamma\gamma')^2\, e^{-H/k_BT_L} \frac{\partial C}{\partial H}\right) \end{aligned} \tag{4.20}$$

This is good news. Inelastic acoustic scattering expression has taken a self-adjoint form which is very convenient for discretization; there are no first derivatives. This formulation, to the best of our knowledge, is new. Original spherical-harmonic work in [20, 32, 33, 47, 48] used the original expression in Eq. (4.15) for discretization.

It may be pointed out that if we use inelastic acoustic phonons, we cannot apply Vecchi's scheme [27–29] to incorporate the full-band band-structure in the spherical-harmonic approach. This is because the expression for inelastic acoustic phonons, Eq. (4.20), has an explicit reference to the dispersion relation $\gamma(\varepsilon)$, while the other scattering terms and the space-dependent term have a reference only to function $g(\varepsilon)$.

4.2.3.4 Ionized Impurity Scattering

Ionized impurity is elastic, it does not change the energy of electrons after scattering. The scattering term for the Y_0^0 term was derived to be zero in Eq. (2.83) of Section 2.5.4, and this remains the same in the new C variable:

$$\left[\frac{\partial F_0^0(H)}{\partial t}\right]_{\text{impurity}} = 0 \qquad \text{Implies} \qquad \left[\frac{\partial C(H)}{\partial t}\right]_{\text{impurity}} = 0 \tag{4.21}$$

4.2.3.5 Impact Ionization Scattering

We start from Eq. (2.94) in Section 2.5.5. It was derived impact ionization in $f_0^0(\varepsilon)$ variable, we re-write in terms of $F_0^0(H)$ variable:

$$\begin{aligned} g(H+q\phi)\left[\frac{\partial F_0^0(H)}{\partial t}\right]_{\text{impact}} &= 4g(2H+2q\phi+\varepsilon_g)\frac{F_0^0(2H+q\phi+\varepsilon_g)}{\tau_{\text{impact}}(2H+2q\phi+\varepsilon_g)} \\ &\quad - g(H+q\phi)\frac{F_0^0(H)}{\tau_{\text{impact}}(H+q\phi)} \end{aligned} \tag{4.22}$$

The first term on the right-hand side corresponds to the scattering *in*, and the second term corresponds to the scattering *out*. By substituting $F_0^0 =$

$C\exp(-H/k_BT_L)$ we get

$$\begin{aligned} g(H+q\phi)\left[\frac{\partial F_0^0(H)}{\partial t}\right]_{\text{impact}} &= g(H+q\phi)\left[\frac{\partial\left(e^{-H/k_BT_L}C(H)\right)}{\partial t}\right]_{\text{impact}} \\ &= 4g(2H+2q\phi+\varepsilon_g)\cdot \\ &\quad \cdot\frac{e^{-(2H+q\phi+\varepsilon_g)/k_BT_L}C(2H+q\phi+\varepsilon_g)}{\tau_{\text{impact}}(2H+2q\phi+\varepsilon_g)} \\ &\quad - g(H+q\phi)\frac{e^{-H/k_BT_L}C(H)}{\tau_{\text{impact}}(H+q\phi)} \end{aligned} \tag{4.23}$$

By using the definition of the collision operator in C variable (Eq. (4.6)):

$$\left[\frac{\partial C}{\partial t}\right]_{\text{impact}} \equiv e^{H/k_BT_L}\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{impact}} \tag{4.24}$$

we can easily derive the expression for the impact ionization scattering in the C variable:

$$\begin{aligned} g(H+q\phi)\left[\frac{\partial C(H)}{\partial t}\right]_{\text{impact}} &= 4g(2H+2q\phi+\varepsilon_g)\cdot \\ &\quad \cdot\frac{e^{-(H+q\phi+\varepsilon_g)/k_BT_L}C(2H+q\phi+\varepsilon_g)}{\tau_{\text{impact}}(2H+2q\phi+\varepsilon_g)} \\ &\quad - g(H+q\phi)\frac{C(H)}{\tau_{\text{impact}}(H+q\phi)} \end{aligned} \tag{4.25}$$

If we compare the new expression, Eq. (4.25) in C variable, with the original equation one, Eq. (4.22) in F_0^0 variable, we see some similarities and some differences. The scattering *in* term in both is similar; the scattering *out* term, on the other hand, in the new expression Eq. (4.25) has an additional factor $\exp(-(H+q\phi+\varepsilon_g)/k_BT_L)$. This factor at room temperature has the value ranging from $\sim 10^{-19}$ down to $\sim 10^{-38}$. This additional factor greatly reduces the magnitude of the off-diagonal coefficient of the impact-ionization expression in C variable.

4.2.4 Final Equation in Pre-Maxwellian Variable

We are now in a position to assemble the pieces together to form our final expression of the spherical-harmonic Boltzmann equation in the new C variable. In Eq. (4.8) we substitute the expressions for optical phonon, Eq. (4.13); inter-band phonon, Eq. (4.14); acoustic phonon, Eq. (4.20); ionized impurity, Eq. (4.21); and impact ionization scattering, Eq. (4.25):

$$\begin{aligned}
&\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial C}{\partial x}\right) && \text{(Space-Dependent)}\\
&+3g\left\{c_{\text{opt}} g^{-} e^{\hbar\omega_{\text{opt}}/k_B T_L}\left(C(H-\hbar\omega_{\text{opt}})-C(H)\right)+\right.\\
&\quad\left. c_{\text{opt}} g^{+}\left(C(H+\hbar\omega_{\text{opt}})-C(H)\right)\right\} && \text{(Optical)}\\
&+3g\left\{c_{\text{ib}} g^{-(\nu')} e^{\hbar\omega_{\text{ib}}/k_B T_L}\left(C^{(\nu')}(H-\hbar\omega_{\text{ib}})-C(H)\right)+\right.\\
&\quad\left. c_{\text{ib}} g^{+(\nu')}\left(C^{(\nu')}(H+\hbar\omega_{\text{ib}})-C(H)\right)\right\} && \text{(Inter-Band)}\\
&+3c_{\text{ac, gC}}^{\text{inelastic}}\, e^{H/k_B T_L}\frac{\partial}{\partial H}\left((\gamma\gamma')^2\, e^{-H/k_B T_L}\frac{\partial C}{\partial H}\right) && \text{(Acoustic)}\\
&+3\left\{4g(2H+2q\phi+\varepsilon_g)e^{-(H+q\phi+\varepsilon_g)/k_B T_L}\right.\\
&\quad\cdot\frac{C(2H+q\phi+\varepsilon_g)}{\tau_{\text{impact}}(2H+2q\phi+\varepsilon_g)}-\\
&\quad\left. g(H+q\phi)\,\frac{C(H)}{\tau_{\text{impact}}(H+q\phi)}\right\} && \text{(Impact-Ionization)}\\
&=0 && (4.26)
\end{aligned}$$

This partial differential-difference equation, along with the boundary conditions, represents the complete boundary-value problem to be solved within a semiconductor device. This equation may be compared with the final form of the Boltzmann equation in the original variable, Eq. (2.100).

4.3 Discretization

It is shown in Appendix A on page 173 that the spherical-harmonic Boltzmann equation is essentially a conservation equation. Its discretization, too, must reflect this conservative property. Therefore, for the spherical-harmonic Boltzmann equation we adopt the current-conserving *control-volume* discretization [113, 114], sometimes also known as box-integration discretization in semiconductor simulation [79, 102–104]. The control-volume approach has a very attractive property: the discretization itself assures the conservation of electron current over any group of control volumes, and, of course, over the whole domain [113, 114]. This conservation is true for any grid, even a coarse grid, not just in the limiting case of large grid number of grid points. The global conservation will ensure, for example, that the total current at the terminals is conserved.

4.3.1 Principles of Control-Volume Discretization

First step is to construct a grid, and then discretize the spherical-harmonic Boltzmann equation on this grid. To ensure the conservation of electron current, we follow some basic principles of control-volume (CV) discretization [113]. While some principles are general in nature, others are specifically tailored to the discretization of the Boltzmann equation.

CV Principle 1 Discretize the domain with non-overlapping control-volumes. This can be done by constructing a grid in Hamiltonian, H, and space, x, such that it covers the whole domain of simulation.

CV Principle 2 Let the grid be orthogonal. This ensures that the fluxes through the control-volume faces are easy to evaluate.

CV Principle 3 Locate the control-volume faces at the mid points of the grid points [113, 114].

CV Principle 4 While the grid is allowed to be non-uniform in space, it is required to be uniform in Hamiltonian. In addition, the spacing in Hamiltonian ΔH should be an integer division of the optical phonon energy: $\Delta H = \hbar\omega_{\mathrm{opt}}/k_{\mathrm{opt}}$, where k_{opt} is a positive integer. This is essential to ensure that optical phonons scatterings conserve electron flux; or the identity of Eq. (A.8) is satisfied by the discretization. Satisfying identity of Eq. (A.8) is crucial to assuring current current-conservation, as can be seen by the proof of current-continuity in Appendix A on page 173.

CV Principle 5 In our band-structure, as seen in Table 3.1 on page 46, the inter-band energy is equal to the optical phonon energy: $\hbar\omega_{\mathrm{ib}} = \hbar\omega_{\mathrm{opt}}$. Therefore, the same grid spacing in Hamiltonian will work for inter-band phonons; identity of Eq. (A.11) will be satisfied.

CV Principle 6 Discretize the equations while accounting for the flux at every control-volume face.

CV Principle 7 If two control volumes share a face, then they must both use same expression for the flux through it. The flux that leaves one control-volume is identical to the flux entering another. This ensures conservation. This principle applies to both space-dependent term as well as the inelastic-acoustic-phonon term. The expression for the flux, in other words, is unique to the face and not to the control-volume.

4.3.2 Discretization of the Boltzmann Equation

Our first task is to construct a grid on the domain of Eq. (4.8), which for a one-dimensional case is $-q\phi(x) \leq H \leq \varepsilon_{\max} - q\phi(x)$ and $L_{\min} \leq x \leq L_{\max}$. A schematic of this orthogonal grid on the domain is shown in Fig. 4.1 on the following page. Keeping the control-volume principles in mind, we construct the grid in Fig. 4.2 on page 65. There are two possibilities to choose from for deciding on the location of the control-volume faces [113, 114]: (a) the control-volume faces located exactly mid-way between grid points; and (b) grid point located exactly at the center of the control-volume. We adopt the first approach.

To derive the discretized equations we start from Eq. (4.8):

$$\frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right) + 3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} = 0 \tag{4.27}$$

We have used a new variable $\kappa = \tau u_g^2 g$. The variable is reminiscent of conductivity, κ, in the classical heat equation [113]:

$$\frac{d}{dx}\left(\kappa\frac{dT}{dx}\right) + \text{Source} = 0 \tag{4.28}$$

The similarity of this heat equation with Eq. (4.27) is obvious. The driving force, gradient of temperature T, is the analogue of the gradient of the distribution. Conductivity κ is analogous with the product $\tau u_g^2 g$. In Fig. 4.2 on page 65 we consider a control volume:

$$\Omega_P = \Big[x_w \leq x \leq x_e\Big] \times \left[\left(k - \frac{1}{2}\right)\Delta H \leq H \leq \left(k + \frac{1}{2}\right)\Delta H\right] \tag{4.29}$$

Integrating Eq. (4.27) in the control volume we get:

$$\iint_{\Omega_P}\left\{\frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right) + 3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}}\right\}\,dx\,dH = 0 \tag{4.30}$$

$$\iint_{\Omega_P}\frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right)\,dx\,dH + \iint_{\Omega_P} 3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}}\,dx\,dH = 0 \tag{4.31}$$

The two terms in the equation can be discretized with a second-order accurate approximation:

$$\begin{aligned}\iint_{\Omega_P}\frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right)\,dx\,dH &= \Delta H\int_x \frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right)\,dx + \mathcal{O}\left((\Delta H)^2\right) \\ &\approx \Delta H\left(\mathcal{F}_e - \mathcal{F}_w\right)\end{aligned} \tag{4.32}$$

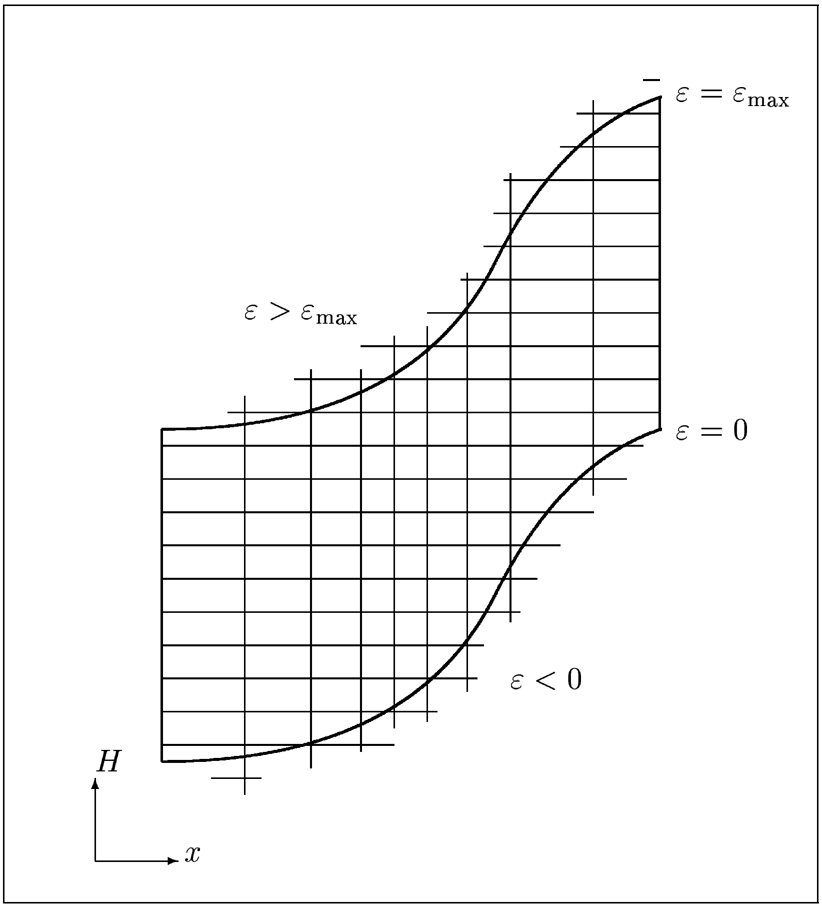

$\varepsilon = \varepsilon_{\text{max}}$
$\varepsilon > \varepsilon_{\text{max}}$
$\varepsilon = 0$
$\varepsilon < 0$
H
x

Figure 4.1: Schematic of the grid on the domain of the spherical-harmonic Boltzmann equation. In a given potential, $\phi(x)$, the domain of the spherical-harmonic differential equation in x-H is $-q\phi(x) \leq H \leq \varepsilon_{\mathsf{max}} - q\phi(x)$ and $L_{\mathsf{min}} \leq x \leq L_{\mathsf{max}}$. The orthogonal grid is uniform in H and non-uniform in x. The intersections of the grid lines are the grid points. Discretization of the Boltzmann is performed only for those grid points that lies within the curvilinear domain.

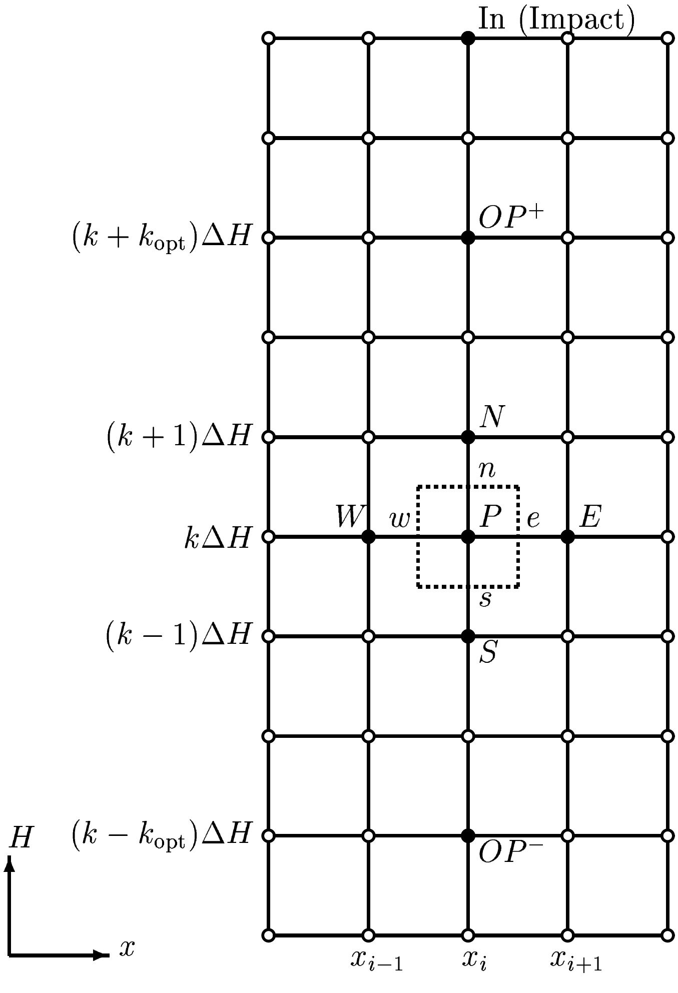

In (Impact)
$(k+k_{\rm opt})\Delta H$
OP^+
N
$(k+1)\Delta H$
n
W
w
P
e
E
$k\Delta H$
s
$(k-1)\Delta H$
S
H
$(k-k_{\rm opt})\Delta H$
OP^-
x
x_{i-1}
x_i
x_{i+1}

Figure 4.2: Grid for the solution of the spherical-harmonic Boltzmann equation. This is a orthogonal Cartesian grid in space dimension, x, and Hamiltonian, H. Filled circles indicate the stencil, or the computational molecule used in the discretization. Control-volume is indicated by the dashed line; control-volume faces are located mid-way between grid points. Grid spacing in Hamiltonian is uniform with ΔH as the spacing; ΔH is chosen such that $k_{\text{opt}}\Delta H = \hbar\omega_{\text{opt}}$ the optical phonon energy jump (In the figure $k_{\text{opt}} = 3$). Grid spacing in space is non-uniform. Inter-band phonon grid points, $IB^{\pm}$, are not drawn, but they can be visualized to be coincident with optical points, $OP^{\pm}$, except that they are in a different band. The point marked "In (Impact)" is the grid point for the scattering-in part of the impact ionization term.

$$\begin{aligned}\iint_{\Omega_P} 3g \left[\frac{\partial C}{\partial t}\right]_{\text{coll}} dx\, dH &= 3g_P \left[\frac{\partial C}{\partial t}\right]_{\text{coll},P} \Delta H \Delta x_P + \\ &\quad \mathcal{O}\left((\Delta H)^2\right) + \mathcal{O}\left((\Delta x_P)^2\right) \\ &\approx 3g_P \left[\frac{\partial C}{\partial t}\right]_{\text{coll},P} \Delta H \Delta x_P\end{aligned} \tag{4.33}$$

Where $\Delta x_P = x_e - x_w = (x_E - x_W)/2$ where $x_e = (x_i + x_{i+1})/2$, and $x_w = (x_{i-1} + x_i)/2$ and $g_P = g(\text{at point } P)$ and $[\partial C/\partial t]_{\text{coll},P}$ evaluated at point P. The collision term is discretized by assuming that the value at point P prevails in the whole control volume. The fluxes are given by second-order accurate central-difference approximations:

$$\mathcal{F}_e = \kappa_e \left(\frac{C_E - C_P}{\Delta x_e}\right) + \mathcal{O}\left((\Delta x_e)^2\right) \tag{4.34}$$

$$\mathcal{F}_w = \kappa_w \left(\frac{C_P - C_W}{\Delta x_w}\right) + \mathcal{O}\left((\Delta x_w)^2\right) \tag{4.35}$$

Where $\Delta x_e = x_E - x_P$, and $\Delta x_w = x_P - x_W$, and The "conductivities" at the control-volume face are κ_e and κ_w. Our first guess might have been to approximate the them as an arithmetic mean of the "conductivity" at the grid points: at the point e we would have written $\kappa_e = (\kappa_E + \kappa_P)/2$. In fact, this mean is second-order accurate. Patankar [113, 116], for the heat-transfer problem, proposed in 1978 that a harmonic mean is a better approximation instead. The interface "conductivity," as derived in Appendix C on page 179, is:

$$\kappa_e = \frac{2\,\kappa_E\,\kappa_P}{\kappa_E + \kappa_P} \tag{4.36}$$

$$\kappa_w = \frac{2\,\kappa_W\,\kappa_P}{\kappa_W + \kappa_P} \tag{4.37}$$

If, for instance, $\kappa_E = 0$ then the harmonic-mean formula will give $\kappa_e = 0$; the arithmetic mean, on the other hand, will give a non-zero value. This behavior of the harmonic mean helps in the energy boundary condition: If the grid point E is outside the domain, then $\kappa_E = 0$, which would make $\kappa_e = 0$; this implies that all references to the grid point C_E will vanish in the discretized equations. For further discussion on harmonic-mean scheme, refer to Appendix C on page 179.

Eq. (4.31)—after substituting Eq. (4.34) and Eq. (4.35) in Eq. (4.32), along with Eq. (4.33)—reads

$$\frac{1}{\Delta x_P}\left\{\kappa_e \left(\frac{C_E - C_P}{\Delta x_e}\right) - \kappa_w \left(\frac{C_P - C_W}{\Delta x_w}\right)\right\} + 3g_P \left[\frac{\partial C}{\partial t}\right]_{\text{coll},P} = 0 \tag{4.38}$$

Which, after rearranging, is

$$\frac{1}{\Delta x_P}\left\{\frac{\kappa_e}{\Delta x_e}C_E - \left(\frac{\kappa_e}{\Delta x_e}+\frac{\kappa_w}{\Delta x_w}\right)C_P + \frac{\kappa_w}{\Delta x_w}C_W\right\} + 3g_P\left[\frac{\partial C}{\partial t}\right]_{\text{coll},\,P} = 0 \tag{4.39}$$

This discretization is second-order accurate: the discretization error goes to zero as the second power of the grid spacing.

$$\text{Discretization Error} = \mathcal{O}\left((\Delta x)^2\right) + \mathcal{O}\left((\Delta H)^2\right) \tag{4.40}$$

4.3.3 Discretization of Collision Terms

4.3.3.1 Optical Phonon Scattering

The discretization of optical phonon term, Eq. (4.13), is also straightforward:

$$\begin{aligned} g_P\left[\frac{\partial C}{\partial t}\right]_{\text{opt},P} &= c_{\text{opt}}\, g_P\, g^{-}_{OP}\, e^{\hbar\omega_{\text{opt}}/k_B T_L}\left(C^{-}_{OP} - C_P\right) + \\ &\quad c_{\text{opt}}\, g_P\, g^{+}_{OP}\left(C^{+}_{OP} - C_P\right) \end{aligned} \tag{4.41}$$

Where $g^{\pm}_{OP} = g(\text{at } \pm\hbar\omega_{\text{opt}} \text{ from point } P) = g(\text{at point } OP^{\pm})$; and $C^{+}_{OP} = C(x_i,\, H_k + \hbar\omega_{\text{opt}})$ and $C^{-}_{OP} = C(x_i,\, H_k - \hbar\omega_{\text{opt}})$. Evaluating $C^{\pm}_{OP}$ has become accurate because our grid in Hamiltonian is constructed such that $\Delta H = \hbar\omega_{\text{opt}}/k_{\text{opt}}$, where k_{opt} is a positive integer (Refer to control-volume principle 4 on page 62 on page 62). If that was not the case we would have to do either of the following: (a) interpolate to get a value for $C^{\pm}_{OP}$; or (b) change the value of effective $\hbar\omega_{\text{opt}}$ to $k_{\text{opt}}\,\Delta H$ where $k_{\text{opt}} = \text{Integer}(\hbar\omega_{\text{opt}}/\Delta H + 0.5)$, as was done in [32, 33, 48]. Both the methods introduce errors: Interpolating, specifically, will prevent us from satisfying the identity of Eq. (A.8), which is crucial to maintaining current continuity in the discretized equations.

4.3.3.2 Inter-Band Optical Phonon Scattering

Expression for inter-band optical phonons, Eq. (4.14), is similar to that for optical phonons. Since in our band-structure the inter-band optical phonon energy jump is the same as optical phonon energy jump, inter-band optical phonon energy is an integer multiple of grid spacing in Hamiltonian too; therefore, same grid spacing works for both.

$$\begin{aligned} g_P\left[\frac{\partial C}{\partial t}\right]^{(\text{This Band}\leftrightarrow\nu')}_{\text{ib},P} &= c_{\text{ib}}\, g_P\, g^{-(\nu')}_{IB}\, e^{\hbar\omega_{\text{ib}}/k_B T_L}\left(C^{-(\nu')}_{IB} - C_P\right) + \\ &\quad c_{\text{ib}}\, g_P\, g^{+(\nu')}_{IB}\left(C^{+(\nu')}_{IB} - C_P\right) \end{aligned} \tag{4.42}$$

Where $g_{IB}^{\pm(\nu')}$ and $C_{IB}^{\pm(\nu')}$ is evaluated in the other bands to which we are scattering. (Points $IB^{\pm}$ are not shown in Fig. 4.2 on page 65; the reader may imagine these points to be coincident with points $OP^{\pm}$, except that they are in a different band.)

4.3.3.3 Inelastic Acoustic Phonon Scattering

Inelastic acoustic phonons expression, Eq. (4.20), is discretized at point P as:

$$\begin{aligned} g_P \left[\frac{\partial C}{\partial t}\right]_{\text{ac}, P} &= c_{\text{ac, gC}}^{\text{inelastic}} \exp(H_P/k_B T_L) \frac{1}{\Delta H} \cdot \\ &\quad \left\{ (\gamma\gamma')_n^2 \exp(-H_n/k_B T_L) \frac{C_N - C_P}{\Delta H} \right. \\ &\quad \left. -(\gamma\gamma')_s^2 \exp(-H_s/k_B T_L) \frac{C_P - C_S}{\Delta H} \right\} \\ &= c_{\text{ac, gC}}^{\text{inelastic}} \frac{1}{\Delta H^2} \left\{ (\gamma\gamma')_n^2 \exp(-\Delta H/2k_B T_L)\ (C_N - C_P) \right. \\ &\quad \left. -(\gamma\gamma')_s^2 \exp(\Delta H/2k_B T_L)\ (C_P - C_S) \right\} \end{aligned} \tag{4.43}$$

Where the subscript n and s, as shown in Fig. 4.2 on page 65, refer to the control-volume face to the north and south. $H_P = H_k$, $H_n = H_k + \Delta H/2$, and $H_s = H_k - \Delta H/2$.

4.3.3.4 Impact Ionization Scattering

The expression for impact ionization, Eq. (4.25), is easy to discretize:

$$\begin{aligned} g_P \left[\frac{\partial C(H)}{\partial t}\right]_{\text{impact}, P} &= 4 g_{\text{impact}, P}^{\text{in}} \cdot \frac{e^{-\Delta H_P^{\text{impact}}/k_B T_L} C_{\text{impact}, P}^{\text{in}}}{\tau_{\text{impact}, P}^{\text{in}}} \\ &\quad - g_P \frac{C_P}{\tau_{\text{impact}, P}} \end{aligned} \tag{4.44}$$

where $\tau_{\text{impact}, P} = \tau_{\text{impact}}(H_P + q\phi_P)$; $\Delta H_P^{\text{impact}} = H_P + q\phi_P + \varepsilon_g$; $C_{\text{impact}, P}^{\text{in}} = C(2H_P + q\phi_P + \varepsilon_g)$ is evaluated at the closest grid point; $g_{\text{impact}, P}^{\text{in}} = g(2H_P + 2q\phi_P + \varepsilon_g)$ is also evaluated at that grid point; and $\tau_{\text{impact}, P}^{\text{in}} = \tau_{\text{impact}}(2H_P + 2q\phi_P + \varepsilon_g)$ too is evaluated at the closest grid point. In Fig. 4.2 on page 65 the grid point marked "In (Impact)" is the one used for the incoming electrons.

4.3.4 Putting it Together

By substituting the discretized scattering terms from Eq. (4.41), Eq. (4.42), Eq. (4.43), and Eq. (4.44) into Eq. (4.39) we get the discretized version of the spherical-harmonic Boltzmann equation:

$$\begin{aligned} A_E C_E + A_W C_W + A_N C_N + A_S C_S + \\ A^+_{OP} C^+_{OP} + A^-_{OP} C^-_{OP} + A^+_{IB} C^+_{IB} + A^-_{IB} C^-_{IB} + A^{\text{in}}_{\text{impact}} C^{\text{in}}_{\text{impact}} \\ - \left(A_E + A_W + A_N + A_S + A^+_{OP} + A^-_{OP} + A^+_{IB} + A^-_{IB} + \right. \\ \left. A_{\text{impact}} \right) C_P = 0 \end{aligned} \tag{4.45}$$

where the coefficients are as follows:

$$A_E = \frac{\kappa_e}{\Delta x_P \Delta x_e} \tag{4.46}$$

$$A_W = \frac{\kappa_w}{\Delta x_P \Delta x_w} \tag{4.47}$$

$$A_N = \frac{3 c^{\text{inelastic}}_{\text{ac, gC}} (\gamma\gamma')^2_n \, e^{-\Delta H / 2 k_B T_L}}{\Delta H^2} \tag{4.48}$$

$$A_S = \frac{3 c^{\text{inelastic}}_{\text{ac, gC}} (\gamma\gamma')^2_s \, e^{\Delta H / 2 k_B T_L}}{\Delta H^2} \tag{4.49}$$

$$A^+_{OP} = 3 c_{\text{opt}} \, g_P \, g^+_{OP} \tag{4.50}$$

$$A^-_{OP} = 3 c_{\text{opt}} \, g_P \, g^-_{OP} \, e^{\hbar\omega_{\text{opt}} / k_B T_L} \tag{4.51}$$

$$A^+_{IB} = 3 c_{\text{ib}} \, g_P \, g^{+(\nu')}_{IB} \tag{4.52}$$

$$A^-_{IB} = 3 c_{\text{ib}} \, g_P \, g^{-(\nu')}_{IB} \, e^{\hbar\omega_{\text{ib}} / k_B T_L} \tag{4.53}$$

$$A^{\text{in}}_{\text{impact}} = 3 \times 4 g^{\text{in}}_{\text{impact}, P} \frac{e^{-\Delta H^{impact}_P / k_B T_L}}{\tau^{\text{in}}_{\text{impact}, P}} \tag{4.54}$$

$$A_{\text{impact}} = 3 \frac{g_P}{\tau_{\text{impact}, P}} \tag{4.55}$$

The discretized Boltzmann equation Eq. (4.45), along with the boundary condition equations, can be written as a matrix equation:

$$\bar{\bar{A}} \bar{C} = \bar{Q} \tag{4.56}$$

The coefficients form the matrix $\bar{\bar{A}}$; unknowns are in the vector $\bar{C}$; and the right-hand-side vector $\bar{Q}$ is the known part of the equation.

4.3.5 Boundary Conditions

Most publications in device simulation concentrate on the description of discretization at the interior grid points, but treat discretization at boundary a little cavalierly (p. 189 in [79]). As a redress for this lament, we will discuss the implementation of the boundary conditions in this sub-section. The boundary conditions developed in Section 2.7 on page 35 for the ohmic, insulator, and the energy boundary are discussed below.

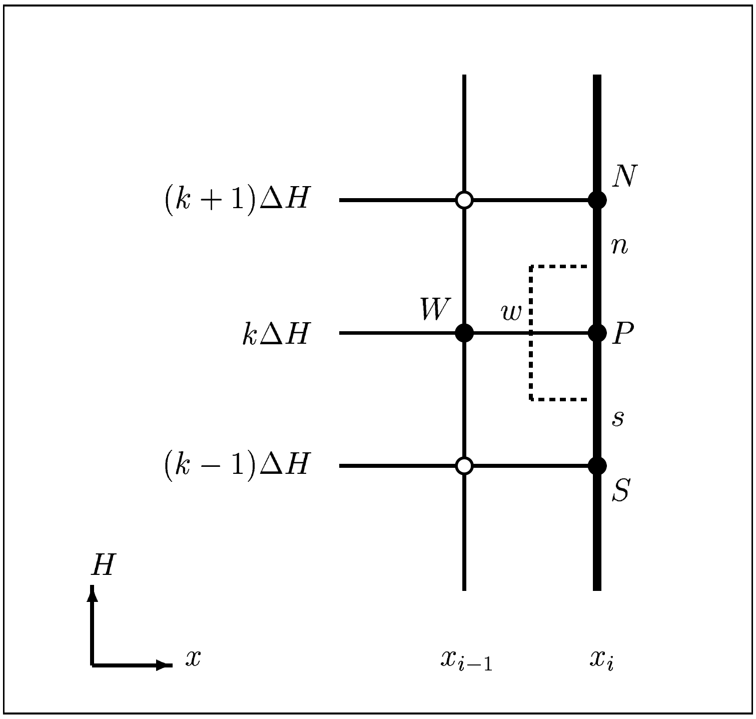


Figure 4.3: The control-volume at an insulator or ohmic boundary. The half control-volume is half of its normal size, and has a grid point at the control-volume face. The heavy line indicates a interface contact or an artificial boundary. The interface contact could be an ohmic contact or an insulator. For the Dirichlet boundary condition we set unknown to the known value. For Neumann boundary condition we can set the flux at the "eastern" side to the specified value.

4.3.5.1 Ohmic and Insulator Boundary Condition

At a material interface or artificial boundary, as shown in Fig. 4.3, the control-volumes is half of its normal size and has a grid point on its face. The ohmic boundary condition, as discussed in Section 2.7.1, is a Dirichlet condition. This poses no problem; we simply set the grid variable to a known value. At an insulating boundary, as discussed in Section 2.7.2, the current flux is zero. This is easily set in the code by setting the "conductivity" κ on that control-volume face to zero. In Fig. 4.3, for instance, we would set $\kappa_e = 0$; other fluxes would be treated normally.

4.3.5.2 Energy Boundary Condition

The application of the energy boundary-condition is slightly less straightforward. In Section 2.7.3 the energy boundary condition was derived as Eq. (2.113). One approach might be to simply discretize this equation. But this is not necessary. We recall from that section that the boundary condition in energy was derived from the spherical-harmonic Boltzmann equation itself by setting $\varepsilon \to 0$ or $\varepsilon \to \varepsilon_{\max}$. Therefore, we can set $\varepsilon \to 0$ or $\varepsilon \to \varepsilon_{\max}$ in the discretized equation itself. This makes the code simple, and assures that all fluxes are accounted consistently.

As $\varepsilon \to 0$ or $\varepsilon \to \varepsilon_{\max}$ the "conductivity" $\kappa = \tau u_g^2 g \to 0$. The harmonic-mean formula, Eq. (4.36) or Eq. (4.37), would result in the interface "conductivity" becoming zero as well. The interface "conductivity" going to zero would ensure that the flux associated with that face is zero, and any reference to the unknown C at boundary is removed from the final discretized equation. For example, let us consider the case shown in Fig. 4.4 on the next page. Since the grid point E is outside the domain, we set $\kappa_E = 0$. By the harmonic-mean formula of Eq. (4.36) we get $\kappa_e = \kappa_E \cdot \kappa_P/(\kappa_E + \kappa_P) = 0$. This sets the flux on the face "e" to zero; which is the intent of the energy boundary condition. When $\kappa_e = 0$, all references to the grid point C_E will be killed in the discretized equation Eq. (4.39). Thus, the boundary condition is realized. It may be mentioned that our implementation of the boundary condition using harmonic-mean of conductivity is similar to what was proposed by Patankar [116] for fluid-flow in domain of arbitrary shape.

For the discretized inelastic acoustic phonons at grid point P in Fig. 4.4, we set $(\gamma\gamma')_s^2 = 0$ in Eq. (4.43). This sets the flux on control-volume face "s" to zero, as well as removes any reference to grid point C_S.

4.4 Advantages of Pre-Maxwellian Variable

Having completed the transformation and discretization of the Boltzmann equation, we come to a very important question: Why transform the Boltzmann equation to the pre-Maxwellian variable? There are four main advantages of using this variable:

1. By rewriting the spherical-harmonic Boltzmann equation in pre-Maxwellian variable, the diagonal dominance of the discretized equations is enhanced. Without this variable transformation, the optical, inter-band optical, and inelastic acoustic phonons spoil the diagonal dominance of the discretized equations; *with* the transformation, they do not. For impact ionization, the diagonal dominance is not restored, but is improved tremendously: the off-diagonal impact-ionization coefficient is drastically reduced, thus enhancing diagonal dominance.
2. The pre-Maxwellian variables account for the rapid exponential variation of the distribution function in *both* energy and space.

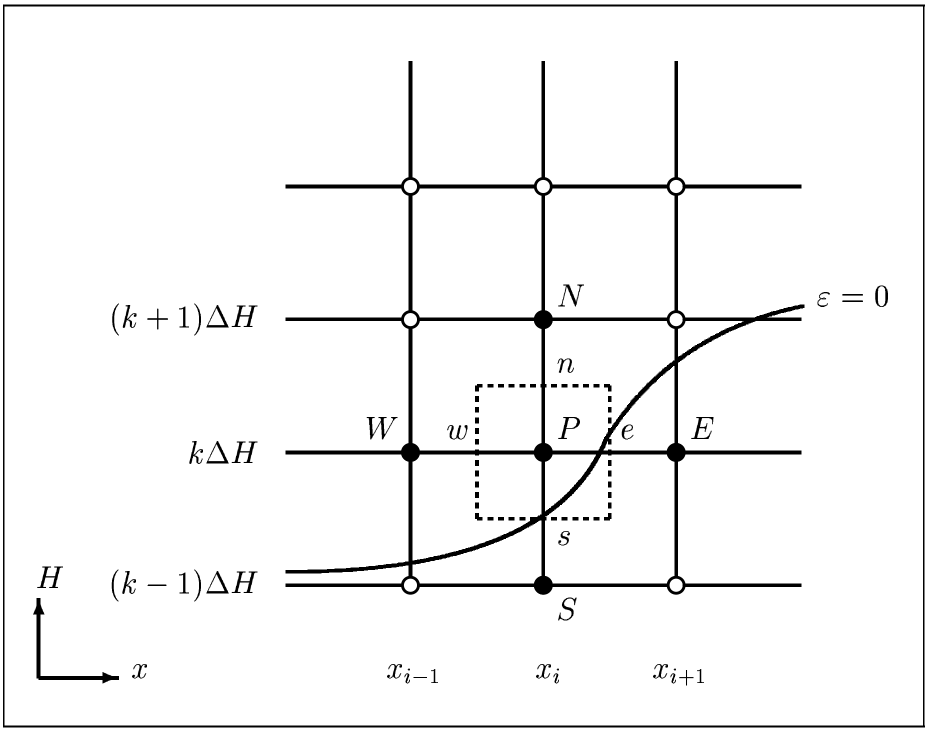


Figure 4.4: Application of the energy boundary condition on the grid. The domain of the equation is marked by the curved line which represents the curve $H = \varepsilon - q\phi(x)$, with $\varepsilon = 0$ or $\varepsilon = \varepsilon_{\mathsf{max}}$. For energy $\varepsilon = 0$ the boundary is shown in the figure. Point E and S are outside the domain of the equation ($\varepsilon < 0$). Although the inner control-volume faces, n and w, are at the mid point of grid points, the boundary control-volume faces, e and s, are assumed to coincide with the domain boundary.

3. The new pre-Maxwellian variable also allow us to write the electron concentrations in Slotboom variables; [98, 102, 104, 105]; and the Poisson equation can be written as a non-linear equation, instead of the usual linear form [70]. this opens the possibility of using superior Poisson-equation solvers, like Mayergoyz' fixed-point algorithm [95–97].

4. These advantages accrue with no penalty in linearity: Boltzmann equation in the new variable is still linear. The structure of the spherical-harmonic Boltzmann equation before and after the transformation is similar; old code, therefore, can be reused with only minor modifications.

Let us discuss some of these advantages in the following sub-sections. For lucidity, the order of discussion has been slightly changed.

4.4.1 Variation of Distribution in Energy and Space

The electron energy distribution function, f, at equilibrium is Maxwellian, that is, it has the form

$$f = \text{Constant} \cdot e^{-\varepsilon/k_B T_L} \tag{4.57}$$

This Maxwellian shape results from the fact that the distribution at equilibrium is dictated by the scattering terms. Mathematically we can show that a Maxwellian distribution is a solution of the optical phonon, acoustic phonon, and inter-band phonons terms equated to zero.

When there are large currents in the presence of strong electric field, electrons are *heated*. The shape of the distribution for heated electrons is non-Maxwellian. But at high energies the scattering rate is large and the high-energy tail may still show a near Maxwellian shape. This means that we can consider the heated distribution to be a perturbation of the Maxwellian case; the constant in Eq. (4.57) is not a constant, but is a function of energy and space. This means that even before we have solved the Boltzmann equation, we have some *a priori* knowledge of the distribution function. We, therefore, define the transformation to new variables $F_0^0(x,H) = C(x,H)\exp(-H/k_B T_L)$. Since $H = \varepsilon - q\phi(x)$, the exponent term has the Maxwellian factor. F_0^0 is also shown depending on $\exp(q\phi/k_B T_L)$. This also makes physical sense as the electron concentration has an exponential dependence on the potential too. In one swoop we have explicitly expressed the energy- and potential-dependence of the distribution:

$$F_0^0(x,H) = C(x,H)e^{-H/k_B T_L} = C(x,H) \underbrace{e^{-\varepsilon/k_B T_L}}_{\text{Energy Dependence}} \underbrace{e^{q\phi(x)/k_B T_L}}_{\text{Space Dependence}} \tag{4.58}$$

4.4.2 Diagonal Dominance in Pre-Maxwellian Variable

A matrix $\bar{\bar{A}}$ whose diagonal element A_P is different in sign than the off-diagonal elements A_l, and satisfies the property

$$|A_P| \geq \sum_l |A_l| \qquad \text{For all equations} \tag{4.59}$$

$$|A_P| > \sum_l |A_l| \qquad \text{For at least one equation} \tag{4.60}$$

is said to be *diagonally-dominant*. For matrix solvers, this is a desirable quality for the solution of matrix equations: direct methods, like Gauss elimination, do not need pivoting; and some iterative methods, like Gauss-Seidel, are assured of convergence [113].

We claim that diagonal-dominance of the discretized Boltzmann equation is enhanced if it is cast in pre-Maxwellian variables. To verify this claim let us examine the Boltzmann equation in original, as well as, new variables.

Comment on Notation Diagonal dominance is a property of a system of equations; it is not property of one equation, or one term in it. But for the sake of discussion below we will refer to *a term of the equation being diagonally dominant.* This should be interpreted as follows: If this term was the only term in the differential equation, and we discretized it, we would get a diagonally-dominant matrix. Therefore, a *diagonally-dominant term*, in our notation, would be one which contributes equally, or more, to the diagonal than it does to the off-diagonal. And a *non-diagonally-dominant term* is the one which contributes more to off-diagonal coefficients than to the diagonal one.

4.4.2.1 Original Variable

When the spherical-harmonic Boltzmann equation in the original form is discretized the equations do not come out to be diagonally dominant. To show this, let us examine the discretization of various components of Eq. (4.3), which reads

$$\frac{\partial}{\partial x}\left(\kappa\frac{\partial F_0^0}{\partial x}\right) + 3g\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{4.61}$$

The space-dependent part of the equation can be discretized easily:

$$\frac{\partial}{\partial x}\left(\kappa\frac{\partial F_0^0}{\partial x}\right) \approx \frac{1}{\Delta x_P}\left\{\frac{\kappa_w}{\Delta x_w}C_W - \left(\frac{\kappa_w}{\Delta x_w}+\frac{\kappa_e}{\Delta x_e}\right)C_P + \frac{\kappa_e}{\Delta x_e}C_E\right\} \tag{4.62}$$

Which can be seen to be diagonally dominant (contributes equally to the diagonal as well as off-diagonal terms).

We Turn our attention to the collision term. Optical phonon collision term, described by Eq. (2.68), term can be discretized at point P to read

$$\begin{aligned} g_P\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{opt},P} = {} & c_{\text{opt}}\, g_P\, g_{OP}^{-}\ \left(F_{OP}^{-} - e^{\hbar\omega_{\text{opt}}/k_BT_L}F_P\right) + \\ & c_{\text{opt}}\, g_P\, g_{OP}^{+}\ \left(e^{\hbar\omega_{\text{opt}}/k_BT_L}F_{OP}^{+} - F_P\right) \end{aligned} \tag{4.63}$$

where $F_{OP}^{-} = F_0^0(H - \hbar\omega_{\text{opt}})$ $F_{OP}^{+} = F_0^0(H + \hbar\omega_{\text{opt}})$. The first term in parenthesis involves the optical phonon absorption and emission between energy ε_P and $\varepsilon_P - \hbar\omega_{\text{opt}}$; we can consider the symbol $\varepsilon_P = H + q\phi$ to be the energy at point P. When the energy ε_P is less than $\hbar\omega_{\text{opt}}$ this term is absent, and the second term remains. The second term deals with emission and absorption between energies ε_P and $\varepsilon_P + \hbar\omega_{\text{opt}}$. The coefficient of the off-diagonal unknown in the second term is larger than the coefficient for the diagonal unknown by a factor of $\exp(\hbar\omega_{\text{opt}}/k_BT_L)$, which is approximately seven for the given $\hbar\omega_{\text{opt}}$ in our band structure at room temperature.

Even at energies larger than $\hbar\omega_{\mathrm{opt}}$, diagonal dominance is not readily achieved. Eq. (4.63) after re-arrangement reads

$$g_P \left[\frac{\partial F_0^0}{\partial t}\right]_{\mathrm{opt},P} = c_{\mathrm{opt}}\, g_P \left\{ g_{OP}^{-}\, F_{OP}^{-} - \left(g_{OP}^{-}\, e^{\hbar\omega_{\mathrm{opt}}/k_B T_L} + g_{OP}^{+}\right) F_P + g_{OP}^{+}\, e^{\hbar\omega_{\mathrm{opt}}/k_B T_L}\, F_{OP}^{+} \right\} \tag{4.64}$$

To have diagonal dominance the diagonal coefficient must be at least as large as the sum of off-diagonal coefficients: $\left(g_{OP}^{-} \exp(\hbar\omega_{\mathrm{opt}}/k_B T_L) + g_{OP}^{+}\right) \geq g_{OP}^{-} + g_{OP}^{+} \exp(\hbar\omega_{\mathrm{opt}}/k_B T_L)$. This is true only if $g_{OP}^{-} \geq g_{OP}^{+}$, that is, the function $g(\varepsilon)$ is decreasing with energy, ε. The function $g(\varepsilon)$ is proportional to the density of states, as seen in Eq. (2.47); the density of states, as seen in Fig. 3.2 on page 47, increases with energy from 0 to 1.75eV, and then again from 2.6 to 2.9eV. Optical phonon does not give diagonally dominant discretization in these energy ranges. This means that the optical phonon scattering term in the original variable will not be diagonally dominant over a large energy range. Same conclusions apply to the inter-band optical phonon term, since it is similar in form to the optical phonon term.

Collision term for the inelastic acoustic phonons, Eq. (4.19) in F_0^0 variable, can discretized in the usual manner. It can be easily shown that such a discretization is diagonally dominant if and only if $(\gamma\gamma')^2$ is decreasing with energy. Say if $\gamma(\varepsilon) = \varepsilon + \alpha\varepsilon^2$, then $(\gamma\gamma')^2 = (\varepsilon + \alpha\varepsilon^2)^2(1 + 2\alpha\varepsilon)^2$ is not decreasing with energy. The discretized inelastic acoustic phonon equation, therefore, is not diagonally dominant either.

Impact Ionization, Eq. (4.22) or Eq. (2.94), is seen to be

$$\begin{aligned} g(H + q\phi) \left[\frac{\partial F_0^0(H)}{\partial t}\right]_{\mathrm{impact}} &= 4g(2H + 2q\phi + \varepsilon_g) \frac{F_0^0(2H + q\phi + \varepsilon_g)}{\tau_{\mathrm{impact}}(2H + 2q\phi + \varepsilon_g)} \\ &\quad - g(H + q\phi) \frac{F_0^0(H)}{\tau_{\mathrm{impact}}(H + q\phi)} \end{aligned} \tag{4.65}$$

The second term is the scattering-out, and the first one is the scattering-in term. The scattering-out term involves the diagonal unknown; and the scattering-in term involves an off-diagonal unknown. At energies below the threshold energy of impact ionization ($\approx$ band gap of silicon, 1.1eV), scattering-out term is zero, while the scattering-in term is not. The impact ionization term, then, has an off-diagonal entry, which hurts diagonal dominance.

At energy larger than the threshold energy of impact ionization, both scattering-in and scattering-out terms are present. The impact ionization term would discretize to a diagonally-dominant equation only if $g(H+q\phi)/\tau_{\mathrm{impact}}(H+$

$q\phi) \geq 4g(2H + 2q\phi + \varepsilon_g)/\tau_{\text{impact}}(2H + 2q\phi + \varepsilon_g)$. This inequality is clearly violated if both the density of states ($\propto g$) and impact-ionization scattering rate ($1/\tau_{\text{impact}}$) increase with energy.

4.4.2.2 Pre-Maxwellian Variable

Now let us turn our attention to the discretization of the spherical-harmonic Boltzmann equation in the pre-Maxwellian variable. We start by reproducing below the spherical-harmonic Boltzmann equation in the pre-Maxwellian variable, Eq. (4.27):

$$\frac{\partial}{\partial x}\left(\kappa \frac{\partial C}{\partial x}\right) + 3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} = 0 \tag{4.66}$$

The first term was discretized using control-volume formulation in Eq. (4.39). This discretization is diagonally dominant—sum of the off-diagonal coefficients is equal to the diagonal one. This is similar to discretization of the space-dependent term with the original distribution-function variable, as in Eq. (4.62).

The diagonal dominance of the optical phonon scattering term is shown below. The discretized optical phonon term, Eq. (4.41), reads

$$\begin{aligned} g_P \left[\frac{\partial C}{\partial t}\right]_{\text{opt},P} = {} & c_{\text{opt}}\, g_P\, g_{OP}^{-}\, e^{\hbar\omega_{\text{opt}}/k_B T_L} \left(C_{OP}^{-} - C_P\right) + \\ & c_{\text{opt}}\, g_P\, g_{OP}^{+} \left(C_{OP}^{+} - C_P\right) \end{aligned} \tag{4.67}$$

On comparing with the optical phonon discretization for normal F_0^0 variable, Eq. (4.63), we see that this expression is diagonally dominant for at *all* energies. A definite improvement.

The discretized inter-band optical phonon scattering term (Eq. (4.42)) is similar to discretized optical phonon term (Eq. (4.41) or Eq. (4.67)); therefore, same conclusions apply to it: discretization of the inter-band optical phonon term results in diagonally-dominant equations at *all* energies.

It is easy to show that the discretization of the inelastic acoustic phonons, as derived in Eq. (4.43), also results in diagonally-dominant equations at *all* energies.

The impact ionization, however, is still slightly problematic. Impact ionization in C variable was derived in Eq. (4.25), and is reproduced here for reference.

$$\begin{aligned} g(H+q\phi)\left[\frac{\partial C(H)}{\partial t}\right]_{\text{impact}} &= 4g(2H+2q\phi+\varepsilon_g)\cdot \\ &\cdot\frac{e^{-(H+q\phi+\varepsilon_g)/k_BT_L}C(2H+q\phi+\varepsilon_g)}{\tau_{\text{impact}}(2H+2q\phi+\varepsilon_g)} \\ &-g(H+q\phi)\frac{C(H)}{\tau_{\text{impact}}(H+q\phi)} \end{aligned} \tag{4.68}$$

Similar to Eq. (4.65), the second term is the scattering-out part, and the first term is the scattering-in part. Further comparing to Eq. (4.65) we see that while the scattering-in term has the same form, the the scattering-out term is different: it has an additional exponential factor of $\exp(-(H+q\phi+\varepsilon_g)/k_BT_L)$. This factor is very small; in Silicon at room temperature it ranges from $\sim 10^{-19}$ down to $\sim 10^{-38}$. It reduces the off-diagonal coefficient by multiplying it, this is very beneficial for diagonal dominance. Even though impact ionization is still not diagonally dominant, situation has been drastically improved—off-diagonal coefficient is much smaller than before.

4.4.2.3 Discussion

With regard to the diagonal dominance, we can summarize our observations:

1. The space-dependent term in the new, as well as the old, formulation is diagonally dominant.
2. The spoilers of diagonal dominance are the collision terms: optical, inter-band, and inelastic acoustic phonon terms; and the impact ionization term.
3. Using the pre-Maxwellian variable, C, makes the optical, inter-band, and inelastic acoustic phonon terms diagonally dominant.
4. The impact-ionization term still does not become diagonally dominant; the scattering-in term spoils this.
5. Even though the impact-ionization term does not become diagonally dominant, using the pre-Maxwellian variable makes the situation much better—it reduces the off-diagonal scattering-in coefficient by a very large exponential factor.

To conclude, we can make the following statement:

> If we rewrite the spherical-harmonic Boltzmann equation in pre-Maxwellian variable and include scattering mechanism of optical, inter-band, acoustic, impurity, and scattering-out impact ionization, then the discretized equations are diagonally dominant.

Term in Equation	**Diagonal Dominance**	
	Distribution Variable	**Pre-Maxwellian Variable**
Space-Dependent	Yes	Yes
Optical Phonon	No	Yes
Inter-band Phonon	No	Yes
Inelastic Acoustic Phonon	No	Yes
Impact Ionization	No	No

Table 4.1: Diagonal dominance of the discretized Boltzmann equation when written in the two variables: the original distribution-function variable, F_0^0; and the new pre-Maxwellian variable, C. Pre-Maxwellian variables make all, but the impact-ionization, terms diagonally dominant. Even though the impact-ionization term does not become diagonally dominant, its off-diagonal coefficient is reduced by a large exponential factor. Since ionized-impurity scattering is elastic, it corresponding term is zero; therefore, it is not mentioned here.

Table 4.1 summarizes our discussion by comparing the the diagonal dominance of the discretized equations with the original variable, and the new pre-Maxwellian variable.

4.4.2.4 Alternate Notation

We can express the conclusions in a more formal fashion. If the coefficient matrix of the discretized pre-Maxwellian-variable Boltzmann equation of Eq. (4.56) is split into the various contributing elements

$$\bar{\bar{A}} = \bar{\bar{A}}_{\text{Elliptic}} + \bar{\bar{A}}_{\text{opt}} + \bar{\bar{A}}_{\text{ib}} + \bar{\bar{A}}_{\text{ac}} + \bar{\bar{A}}_{\text{impact}} \tag{4.69}$$

and the boundary conditions are conveniently divided into the various matrices, then we can say the following: all matrices, except the matrix $\bar{\bar{A}}_{\text{impact}}$, are diagonally dominant. But if we expanded the matrix for the discretized Boltzmann equation in original variable, the only matrix that would be diagonally dominant would be the $\bar{\bar{A}}_{\text{Elliptic}}$ matrix.

4.4.2.5 Why Pre-Maxwellian Variables Give Diagonal Dominance?

Maxwellian distribution ($f_0^0 = C_0 \exp(-\varepsilon/k_B T_L)$, where $C_0 = \text{Constant}$) is the non-trivial solution of the equation

$$\left[\frac{\partial f_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{4.70}$$

if the collision term is optical, inter-band optical, or inelastic acoustic phonon. This can be easily verified by substituting $f_0^0 = C_0 \exp(-\varepsilon/k_B T_L)$ in the following equations: Eq. (2.68) for optical phonons; Eq. (2.72) for inter-band optical phonons; and Eq. (2.77) for inelastic acoustic phonons. For these collision terms, therefore, $C = C_0$ (Constant) satisfies the equation

$$\left[\frac{\partial C}{\partial t}\right]_{\text{collision}} = 0 \tag{4.71}$$

If the discretization of the above equation is

$$\sum_l A_l C_l - A_P C_P = 0 \tag{4.72}$$

then $C = C_0$ (Constant) must also satisfy this discretized equation (Eq. (4.72)). This is the case if and only if

$$A_P = \sum_l A_l \tag{4.73}$$

This equation is a statement of diagonal-dominance. Thus, we have shown that diagonal dominance emerges in pre-Maxwellian variables because the Maxwellian distribution satisfies the collision-term-equated-to-zero equation.

This explanation has some interesting consequences: (a) We did not see diagonal dominance in the original distribution-function variable because $f_0^0 =$ Constant is not a solution of the equation Eq. (4.70). (b) Impact ionization did not give diagonal dominance because $[\partial f_0^0/\partial t]_{\text{impact}} = 0$ admits only the trivial solution $f_0^0 = 0$. (c) In addition, this insight allows us to predict that future scattering mechanisms which have Maxwellian as the solution of Eq. (4.70) will also give diagonal dominance when transformed to pre-Maxwellian variables. (d) It may be pointed out that this is an interesting situation where physics and numerics are related.

4.4.3 Poisson-Equation Solvers with Pre-Maxwellian Variable

Electron concentration with pre-Maxwellian variable, from Eq. (2.45) and Eq. (2.46), can be expressed in terms of the Slotboom variable $u(x)$ [98, 102, 104, 105]:

$$\begin{aligned} n(x) &= \frac{1}{4\pi^3}\frac{1}{\sqrt{4\pi}}\int_H g\,F_0^0(x,H)\,dH \\ &= \frac{1}{4\pi^3}\frac{1}{\sqrt{4\pi}}\int_H g\,C(x,H)e^{-H/k_BT_L}\,dH \\ &= \frac{1}{4\pi^3}\frac{1}{\sqrt{4\pi}}\int_H g\,C(x,H)e^{-(\varepsilon-q\phi(x))/k_BT_L}\,dH \\ &= \frac{1}{4\pi^3}\frac{1}{\sqrt{4\pi}}\,e^{q\phi(x)/k_BT_L}\int_H g\,C(x,H)e^{-\varepsilon/k_BT_L}\,dH \\ &= u(x)\,e^{q\phi(x)/k_BT_L} \end{aligned} \tag{4.74}$$

This expression of electron concentration can be put to good use by incorporating it in a superior Poisson-equation solver, like Mayergoyz' fixed-point algorithm [95–97]. It is important in this algorithm to express the concentration in Slotboom variables; Poisson equation in this variable then becomes non-linear, as written in Eq. (D.6). In fact, it has been reported that non-linear Poisson equation may be better for self-consistent Monte Carlo simulations [70]. Using pre-Maxwellian variables allows us to express the concentration in Slotboom variables, which makes the Poisson equation non-linear, and thereby permits us to use the fixed-point algorithm.

This fixed-point iteration method has numerous advantages [95]: (a) it is fast; (b) it has global convergence; and (c) it updates the potential by explicit formulas, therefore avoids the solution of matrix equations. Refer to [96, 97] for improvements and comments on this technique. Boltzmann equation in original distribution variable, does not offer this luxury; the choice of Poisson solvers, therefore, is more restrictive.

4.5 Chapter Summary

In this chapter the spherical-harmonic Boltzmann equation was transformed. The distribution function was expressed in a new pre-Maxwellian variable. The inelastic acoustic phonon term was transformed to a novel self-adjoint form. The Boltzmann equation in new variable was discretized by current-conserving control-volume scheme. This chapter then concluded by presenting the advantages of the pre-Maxwellian variables.

CHAPTER 5

SIMULATION OF N$^+$NN$^+$ DEVICE STRUCTURE

5.1 Chapter Introduction

To simulate a semiconductor device we need to self-consistently solve four coupled partial differential equations: two Boltzmann equations, one for the lower bands and another for the upper bands; Poisson equation; and hole-continuity equation. Hole transport, being less important than electron transport, is simulated by the drift-diffusion hole-continuity equation, rather than a rigorous hole Boltzmann equation [62]. Discretization of these equations has been discussed elsewhere in the dissertation—Boltzmann equation discretization was discussed in the previous chapter, and the Poisson and hole-continuity equations are discretized in Appendix D on page 183 and Appendix E on page 189 respectively. This chapter is devoted to developing solution techniques for the discretized equations. The techniques developed and tested on a n$^+$nn$^+$ device in this chapter are used in subsequent chapters to simulate other devices.

The line successive over-relaxation (SOR) matrix solver was found to be an efficient solver for the Boltzmann equations [113, 114]. The four equations are solved self-consistently by decoupled scheme [114], which, in device simulation parlance, is known as the Gummel's decoupled scheme [99, 101, 102]. A one-dimensional sub-micron n$^+$nn$^+$ device structure is simulated by this technique. The choice of the n$^+$nn$^+$ structure is not accidental: It is a common device structure for testing new simulators because (a) it displays non-equilibrium behavior typical of modern devices; (b) it approximates the channel region of a MOSFET in one-dimension; and (c) requires the solution of only one type of carriers. The thermal-equilibrium test case from Appendix B on page 177 was applied on the simulator.

It was observed that the distribution function is susceptible to *numerical noise*—indicative of a poor numerics. The origin of this noise was detected to be an absence of coupling to neighboring points in energy. This was corrected by using acoustic phonons in an inelastic approximation. Although elastic approximation is more common [17, 21, 31, 35–37, 39, 41, 43], the inelastic approximation of Kolodziejczak [59] is not new to spherical-harmonic work [20, 32, 33, 47, 48]. Therefore, the usage of inelastic phonons by itself is not new; what *is* new is that it is identified as the cure for numerical noise. It may be noted that a new discretization for the inelastic acoustic phonons has already been discussed in Chapter 4.

This chapter is organized as follows: Section 5.2 discusses the matrix solution technique for the Boltzmann equation. In Section 5.3 on the facing page the solution of the coupled system of equations by the Gummel method is discussed. The simulator is described in Section 5.4 on page 85, and the results for the n$^+$nn$^+$ device are presented in Section 5.5 on page 85. The thermal-equilibrium test case is applied in Section 5.6 on page 89. Convergence of the Gummel loop and matrix solvers is discussed in Section 5.7 on page 90. Section 5.8 on page 91 discusses the important aspect of noise in the solution and how it is cured.

5.2 Solution of Discretized Equations

There are four coupled differential equations to be solved:

1. Boltzmann equation for band (12) (concatenation of bands (1) & (2))
2. Boltzmann equation for band (34) (concatenation of bands (3) & (4))
3. Poisson Equation
4. Hole-continuity Equation

This section discusses their solution.

The Boltzmann equation was discretized by control-volume formulation in the previous chapter where Eq. (4.45) was derived. The discretized Boltzmann equation can be written in terms of a system of equations of the form

$$\sum_l A_l C_l - A_P C_P = Q_P \tag{5.1}$$

where C can be either $C^{(12)}$ or $C^{(34)}$. Subscript P indicates the point at which the discretized Boltzmann equation is written; summation is over all other points in Eq. (4.45). Q_P is contains the C of the other band due to the interband scattering.

Boltzmann equation is solved by the line successive over-relaxation (SOR) method, as illustrated in Fig. 5.1 on page 84; it can be described as follows:

1. Since all line-SOR matrices are tridiagonal, solve them by the Thomas' tridiagonal matrix algorithm (TDMA) [113,114].
2. Use SOR when updating the unknown: $C^{\mathrm{new}} = \omega C^{\mathrm{tdma}} + (1-\omega)C^{\mathrm{old}}$; where the solution obtained by TDMA is C^{tdma}, and ω is the over-relaxation factor.
3. Use line-SOR in space for points in grid with constant H. Repeat for all possible H.
4. Use line-SOR in energy and group only points connected by the phonon energy jump $\hbar\omega_{\mathrm{opt}}$. Repeat this for all possible groups at a spatial point, and then loop over all space points.
5. Use line-SOR in energy for all points in energy connected by acoustic phonons. Since acoustic phonons at H are connected with points $H \pm \Delta H$, all points at a spatial location are included in one group. Loop over all space points.

The discretization and numerical solution of the Poisson and hole-continuity equations is discussed in Appendix D on page 183 and Appendix E on page 189 respectively.

5.3 Solution of the Coupled Equations

The four differential equations are coupled, that is, the *dominant variable* of one equation occurs in other equations. The coupled system is solved by decoupling the equations: Each equation equation is solved in its dominant variable while treating the other variables as constant, then iterating this through all equations until the whole system is solved [114]. This decoupled scheme is known in device simulation as the Gummel's decoupled method [99,101,102].

Fig. 5.2 on page 86 illustrates this procedure. Initial guesses are generated from a drift-diffusion simulation of the device. The drift-diffusion output provides the initial potential $\phi(x)$, hole concentration, and electron concentration $n(x)$. The distribution function is assumed to be Maxwellian shaped ($C(x,H) = \text{Constant}$) with a electron concentration equal to that provided by the drift-diffusion solution. We then solve the two Boltzmann equations by line-SOR method, followed by a solution the Poisson and hole-continuity equation. Then convergence of the whole system is checked, if it has not converged we iterate over all the equations again. This outer loop, iteration over the equations, is known in semiconductor device simulation as the *Gummel loop* [101].

When solving the Boltzmann equations, the inner loops of line-SOR method need not be taken to full convergence because the other variables are tentative

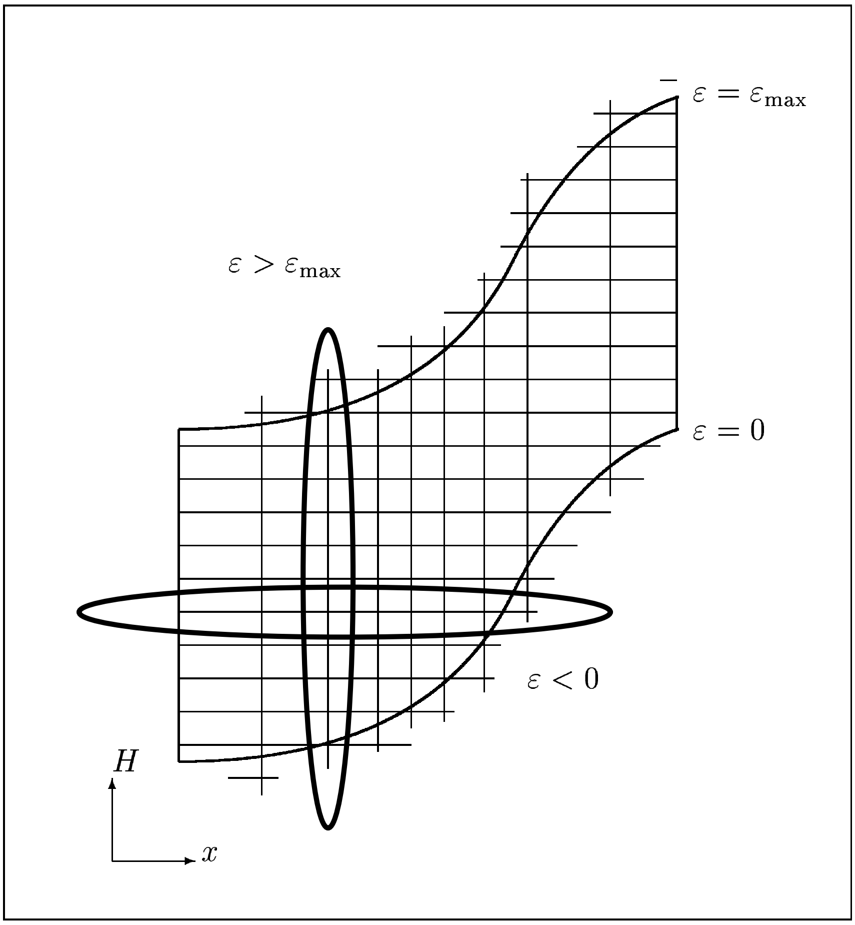

$\varepsilon = \varepsilon_{\max}$
$\varepsilon > \varepsilon_{\max}$
$\varepsilon = 0$
$\varepsilon < 0$
H
x

Figure 5.1: Illustration of the line-SOR approach to solve the Boltzmann equation. Grid points are grouped together in either space or Hamiltonian (energy). Within the group, the matrix is tridiagonal, and is solved by a direct solver. The tridiagonal solver is applied to all grid points in alternating directions. The process is repeated until the iterations converge.

at best [113]. This is achieved in the code by limiting the maximum number of inner loops.

The convergence criterion for the Boltzmann equation is little tricky. One would be tempted to compare $C(x, H)$ from a previous Gummel loop to the new one. This, however, may not be accurate representation of convergence. Every Gummel the electric potential $\phi(x)$ changes. This changes the domain of the Boltzmann equation. A grid point (x, H) no longer represent the same energy in different Gummel loops. The way out of this is to use the relative change of electron concentration between Gummel loops. If m is the current

Gummel loop index, then, the convergence criteria is:

$$\max_{\forall i} \left| \frac{n_i^{(m)} - n_i^{(m-1)}}{n_i^{(m)}} \right| < 10^{-8} \text{ to } 10^{-5} \tag{5.2}$$

To make the convergence test more robust, the code also checks the relative current density between nodes:

$$\max_{\forall i} \left| \frac{J_{i+1/2}^{(m)} - J_{i+1/2}^{(m-1)}}{J_{i+1/2}^{(m)}} \right| < 10^{-8} \text{ to } 10^{-5} \tag{5.3}$$

The Poisson is considered converged when

$$\max_{\forall i} \left| \phi_i^{(m)} - \phi_i^{(m-1)} \right| < \left(10^{-6} \text{ to } 10^{-4}\right) V_t \tag{5.4}$$

where $V_t = k_B T_L / q$, the thermal voltage. Hole-continuity equation is considered converged when

$$\max_{\forall i} \left| \frac{p_i^{(m)} - p_i^{(m-1)}}{p_i^{(m)}} \right| < 10^{-6} \text{ to } 10^{-4} \tag{5.5}$$

5.4 The Simulator

A one-dimensional simulator is written for $\mathrm{n^+nn^+}$ devices and bipolar transistors. It solves the drift-diffusion as an initial guess, followed by a self-consistent solution of the two multi-band Boltzmann equations and the Poisson and hole-continuity equations. It has 2,500 lines of uncommented code in FORTRAN 77, which were executed on a 200 MHz Pentium Pro personal computer (PC) with 64 MB of RAM (Random-Access Memory) running a Linux RedHat 4.0 operating system using the Gnu Fortran compiler. The same simulator is used also used for the one-dimensional BJT of Chapter 6.

5.5 Simulation of a $\mathrm{n^+nn^+}$ Device

In this section the simulation results for a sub-micron $\mathrm{n^+nn^+}$ device are presented and discussed. The $\mathrm{n^+nn^+}$ device structure is a convenient test vehicle to test modern device simulation codes—it displays non-equilibrium behavior typical of sub-micron MOSFET's and Bipolar transistors; requires solution of only one type of carriers; and roughly approximates the channel of a MOSFET unencumbered by two-dimensional effects (p. 265 of [74]).

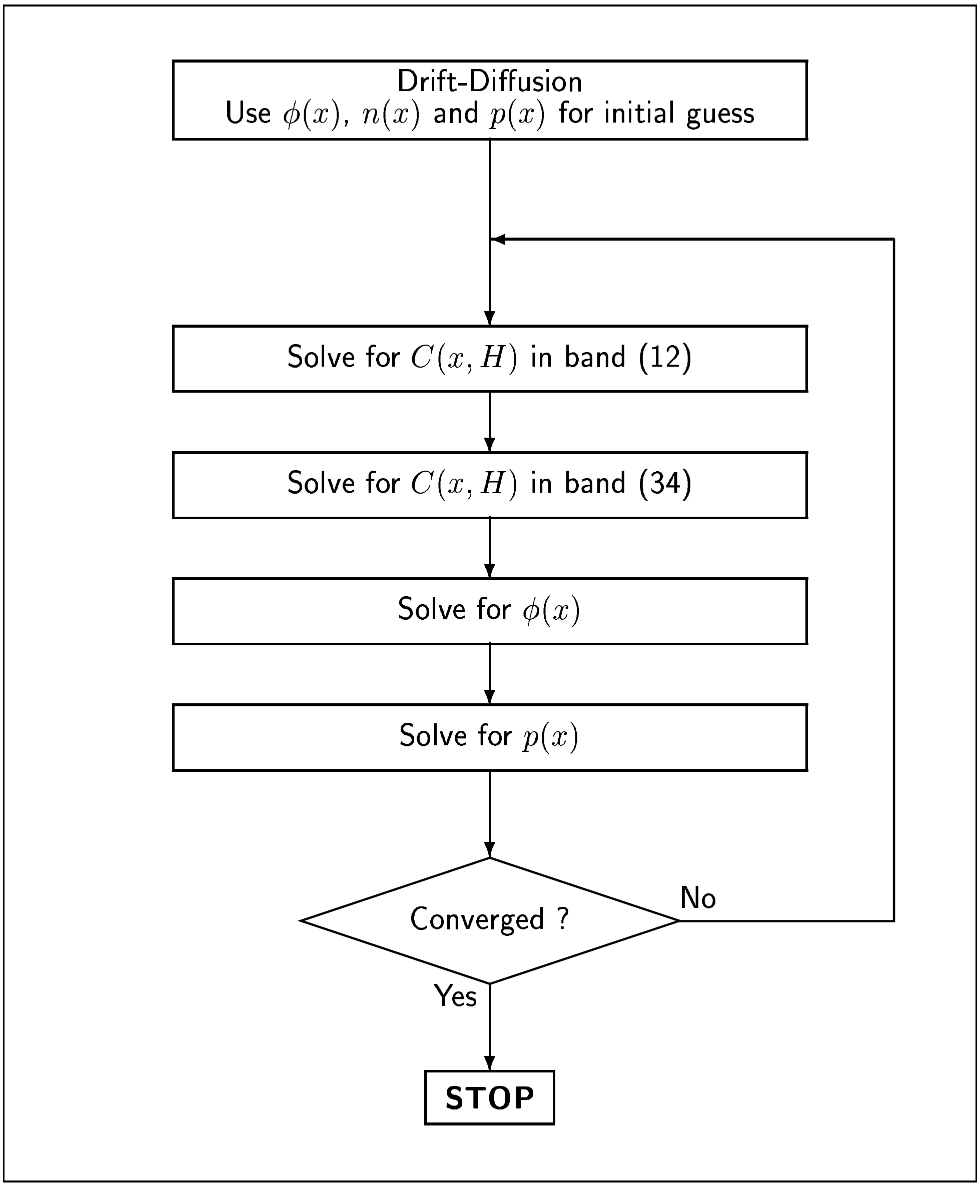

Drift-Diffusion
Use $\phi(x)$, $n(x)$ and $p(x)$ for initial guess
Solve for $C(x,H)$ in band (12)
Solve for $C(x,H)$ in band (34)
Solve for $\phi(x)$
Solve for $p(x)$
Converged ?
No
Yes
STOP

Figure 5.2: The four coupled discretized equations are solved in a decoupled Gummel scheme. Each equation is solved for the dominant variable while assuming the other variables known. After looping through all equations, which is known as the Gummel loop, convergence checks are made to determine whether to terminate the loop or to continue with the iterations. The inner loops of the individual equations are not shown.

5.5.1 Device Structure & Grid

The device structure is shown in Fig. 5.3 on page 94. It has two n$^+$ "source" and "drain" regions of 0.1μm width and doping of 5×10^{17} cm^{-3} sandwiching a lightly-doped n-region of width 0.4μm and doping 2×10^{15} cm^{-3}. There are two ohmic contacts: one at source end, $x = 0$; and the other at drain end, $x = 0.6\mu$m. This device structure, in fact, is quite typical for testing silicon device simulators [109].

The grid is uniform in x and has approximately 60 points. In H the grid is also uniform, the grid spacing is $\Delta H = \hbar\omega_{\text{opt}}/6$, which translates to 349 points in band (12), and 91 points in band (34). The total number grid points is approximately 26,400. For a typical run the code converges in around a hundred Gummel loops and takes about ten wall-clock minutes.

5.5.2 Simulation Results at Various Biases

The ohmic contact at $x = 0$ is grounded, and a bias is applied at the ohmic contact at $x = 0.6\mu$m. This sub-section presents the results for three biases: $V_{\text{applied bias}} = 1$, 3, and 5V. All simulations are at room temperature, $T_L =$ 300K. Fig. 5.4 on page 95 plots the self-consistent electric potential within the device at these biases. There is very little potential drop in the heavily doped n$^+$ regions, most of it occurs in the lightly doped middle region. This is reflected in the electric-field profile within the device as shown in Fig. 5.5 on page 95. Electric field is essentially zero in the heavily doped regions. The maximum occurs at the right n-n$^+$ junction. Near the left n$^+$-n junction we can see the built-in field, which is in the direction opposite to the field in most of the device.

This average velocity resulting from these three biases is shown in Fig. 5.6 on page 96. We can see three regions of interest: (a) low-velocity regions; (b) saturation region; and (c) velocity-overshoot region. In the heavily-doped regions there a large number of electrons, therefore the average velocity is much smaller. As the electrons leave the "source" region to enter into the lightly-doped region (near $\sim 0.1\mu$m) they suddenly experience a large electric field, this causes their velocity to overshoot the saturation velocity. [1] While velocity overshoot can be clearly seen at around 0.2μm for the 5V bias, it is not present for 1V bias because the electric field, and its gradient, is not large enough to cause it. After overshooting the electrons tend to approach the saturation velocity, which in silicon is $\sim 10^7$ cm/s.

Velocity overshoot, as we shall see in later chapters too, occurs in devices whenever electrons enter a region of high electric-field gradient [74,107]. Electrons entering such a region do not have enough time to gain energy and "heat-up"—in other words, they are not in equilibrium with the local electric field. Lower average energy means less scattering and higher mobility and higher velocity. When the electrons do come to an equilibrium, electron velocity saturates. Saturation velocity is a balance between energy gained from electric field, and lost to optical phonons.

---

[1] Velocity overshoot is said to occur when the local average velocity of electrons is larger than what it would have been had the electrons been at equilibrium with the local electric field.

Fig. 5.7 on page 96 shows the electron concentration at the three applied biases. Electron concentration is lower in regions of high velocity, this is because the current, which is $J = nv$, has to be conserved. In fact the current as function of position, which is plotted in Fig. 5.8 on page 97, is constant. This is very satisfying as it validates the current-conserving nature of the control-volume discretization.

Average electron energy is displayed in Fig. 5.9 on page 98. In the "source" region the electrons are near equilibrium and have the average energy of a Maxwellian, $3/2k_BT_L$, which is equal to 0.0388eV at room temperature. As the electrons enter the electric-field region they drift and gain energy. In the "drain" region the average energy relaxes back to its equilibrium value.

The impact-ionization coefficient α_{impact} is the ratio of Generation due to impact ionization G_{impact}, and the current density J:

$$\alpha_{\text{impact}}(x) \triangleq \frac{G_{\text{impact}}(x)}{J(x)} \tag{5.6}$$

Generation rate is calculated in Eq. (2.97) by integrating over the high-energy tail: $G_{\text{impact}}(x) = \int_\varepsilon \tilde{h}(\varepsilon)(f_0^0(\varepsilon, x)/\tau_{\text{impact}}(\varepsilon))\, d\varepsilon$. Impact-ionization coefficient is a measure of impact ionization. For the three biases in the n$^+$nn$^+$ device, we see the impact-ionization coefficient in Fig. 5.10 on page 98. Impact ionization is higher for higher bias because the electrons are heated by the larger electric field. If we compare the location of the maximum of α_{impact} in Fig. 5.10 on page 98 with the location of maximum average energy in Fig. 5.9 on page 98, or the location of maximum electric field in Fig. 5.5 on page 95, we notice that it occurs a little further to the right. This is indicative of the non-equilibrium nature of electron transport—neither electric field, nor average energy can accurately model hot-electron effects like impact-ionization, instead the complete distribution function is required. This phenomenon of delay in impact ionization is also referred to as *dead-space effect* [94], or also as the *dark-space effect* [69].

5.5.3 Drift-Diffusion versus Boltzmann Equation

We have solved the Boltzmann equation self-consistently with the Poisson and the hole-continuity equations. In Fig. 5.11 on page 99 we compare the solution by the self-consistent Boltzmann equation to the solution of drift-diffusion (DD) equations. Near $x = 0.15\mu$m the electron concentration by Boltzmann equation shows a small dip, this is because of the velocity overshoot. Drift-diffusion equations do not show a dip because of the absence of velocity overshoot.

Fig. 5.12 on page 99 compares the electric field obtained by the solution of the self-consistent Boltzmann system with that from a drift-diffusion solution. The differences in electric field are originating directly from the difference in electron concentrations shown in Fig. 5.11 on page 99.

5.5.4 Distribution Function at 5V

Fig. 5.13 on page 100 shows the multi-band distribution function computed by the spherical-harmonic technique. The distribution function at the ohmic contacts is forced by Dirichlet boundary-condition to be Maxwellian ($\propto \exp(-\varepsilon/k_BT_L)$) at lattice temperature. As the electron travel from the "source" region they gain energy from the electric field and the distribution function slowly departs from the Maxwellian shape and *heats-up*. These energetic hot electrons scatter, inter-band scattering, to the upper bands of silicon, as seen by the distribution in band (34). The population of upper bands increases by many orders of magnitude. In the "drain" region electric field is small so the electrons tend to relax back to the Maxwellian shape, until at the ohmic contact at $x = 0.6\mu$m they are forced to be Maxwellian by the Dirichlet boundary-condition. In fact, in the n$^+$ "drain" region one can see two slopes of the distribution, indicative of two populations of electrons coexisting: one which is Maxwellian at room temperature; and the other due to the injection of hot-electrons from the "channel" region [110].

5.6 Thermal Equilibrium

The case of thermal equilibrium in a device is very interesting. There are three motivating factors for the study of the thermal equilibrium. Firstly, this is the only non-trivial case for which the analytical solution of the Boltzmann-Poisson system is available. Hence, it can serve as good check for the model, numerical technique and the computer code. Secondly, since the resulting distribution function is exponential (Maxwellian), this case also represents the worst case scenario for the numerical technique because the solution exhibits boundary-layer phenomena (it is concentrated near zero energy). Thirdly, the solution at thermal equilibrium is interesting in its own right: At thermal equilibrium there can be large electric fields in the device, but there is no current flow (electrons have zero velocity). This can be interpreted as *velocity undershoot* or *velocity damping*: Electron velocity much smaller than what is to be expected from the local electric field.

It is shown in Appendix B on page 177 that at thermal equilibrium the distribution function in a device is Maxwellian, and the potential distribution is given by the thermal-equilibrium non-linear Poisson equation. Our device-simulation computer program was tested for the thermal-equilibrium test case by applying a zero bias to the n$^+$nn$^+$ device. The distribution, as shown in Fig. 5.14 on page 101, is Maxwellian for both the bands; even though there are large electric fields, as seen in Fig. 5.15 on page 102.

5.7 Convergence of Numerical Solution

In this section the convergence behavior of the numerical solution is examined; both for the self-consistent Boltzmann-Poisson system, as well as for the non-self consistent situation.

Fig. 5.16 on page 103 plots the error criteria for all the equations against the outer Gummel loop. The n$^+$nn$^+$ device was biased at 5V. The successive over-relaxation (SOR) factor, ω, was set to 1.4, which was the optimal value as determined by numerical experimentation. The decrease is erratic in the first 80 Gummel loops; after that it decreases smoothly. The erratic behavior can be explained as follows. During the initial Gummel loops the electric potential $\phi(x)$ changes more. The change in $\phi(x)$ causes the domain of the Boltzmann equations to change, causing some grid points to move out, and some grid points to move in. This is a severe change for the numerical scheme and the Boltzmann equations react with large change in electron concentration. When the change in $\phi(x)$ is less, and no grid points are moving-in or moving-out, then the error for the entire system is decreases smoothly. This dependence of the domain on the solution of the Poisson equation has adverse effects in a self-consistent solution. This is an area of potential improvement in the future.

Next we examine the performance of various matrix solvers used for the solution of the discretized Boltzmann equation. From the discussion in Section 5.2 on page 82 four successive over-relaxation (SOR)-type matrix solvers were tested:

1. Line-SOR alternating in space and Hamiltonian (Line-SOR in x & H)
2. Line-SOR in space only (Line-SOR in only x)
3. Line-SOR in Hamiltonian only (Line-SOR in only H)
4. Point-by-point SOR (Point-SOR)

To isolate the convergence of the matrix solver from the rest of the solution, Boltzmann equation in band (12) was solved in a non-self-consistent manner—the potential distribution, which came from the drift-diffusion solution, remained unchanged. The SOR factor ω was set to a convenient value of 1.2.

Since the potential remained unchanged we can now use a different measure of error, relative change in $C^{(12)}(x, H)$:

$$\text{maximum} \ \forall \ i, k \ \left| \frac{C_{i,k}^{(m)} - C_{i,k}^{(m-1)}}{C_{i,k}^{(m)}} \right| .$$

From the Fig. 5.17 on page 104, which plots the convergence behavior of the four solvers, following observations can be made. The convergence of point-SOR technique is painfully slow. This is not surprising; the point-SOR method transmits the boundary-condition information at a rate of one grid interval per iteration [113]. The line-SOR technique would be better as we could bring boundary-condition information to the interior at a faster rate. Two line-SOR possibilities are (a) line-SOR in only H, and (b) and line-SOR in only x are, therefore, faster converging than point SOR, as seen in the figure. We also see that line-SOR in x is more efficient than line-SOR in H. This is easily understandable from the structure of the Boltzmann equation: The Boltzmann equation has coupling in H direction by the optical phonon, equation at H is coupled to equations at $H \pm \hbar\omega_{\text{opt}}$. If the grid spacing in H is $\Delta H = \hbar\omega_{\text{opt}}/n$, then the Boltzmann equation at H naturally couples to, and hence transmits information to, grid points $H \pm n\Delta H$. Since this optical phonon coupling in H transmits information in H direction efficiently, line-SOR in H is not as badly needed as line-SOR in x. Of course, if we still use both, line-SOR in x & H, we get much better convergence rate, as seen in the figure.

5.8 Numerical Noise and Inelastic Acoustic Phonons

5.8.1 Summary of the Idea

In the initial stages of development of the numerics and code for the solution of the Boltzmann equation the computed distribution function often had *numerical noise.* Numerical noise—wiggles or oscillations in the solution—is an artifact of the numerical solution, and it points to a weakness of the numerical technique. Careful investigation of the numerical technique revealed this to be due to the absence of coupling to neighboring points in Hamiltonian. In fact, the spherical-harmonic Boltzmann equation intrinsically does not have a derivative in Hamiltonian—this can seen in Eq. (2.24); or in Eq. (2.100) or Eq. (4.26) without the inelastic-acoustic-phonon term. The prognosis of noise also suggested its cure: Introduction of coupling in the Hamiltonian direction. Coupling was introduced by changing the acoustic phonons from the commonly-used elastic approximation to an inelastic approximation. This hypothesis is corroborated by examining the distribution with elastic acoustic phonons (Fig. 5.18 on page 105) against the one obtained with the inelastic acoustic phonons (Fig. 5.19 on page 106). Next sub-section explains this idea in more detail.

5.8.2 Detailed Explanation

The elastic approximation for acoustic phonons is quite common, and perhaps also justified—the energy of electrons changes very little by acoustic phonon scattering [62,74]. Most spherical-harmonic work, therefore, has also embraced this approximation [17,21,31,35–37,39,41,43]. Elastic approximation makes the acoustic collision term in spherical-harmonic-expanded Boltzmann equation zero, which makes the equation simpler to solve and discretize—no derivative in Hamiltonian to worry about. This, therefore, was also my starting approach, but the computed distribution function was turned out to be noisy, as shown in Fig. 5.18 on page 105.

The distribution function in Fig. 5.18 on page 105 is clearly not correct. To debug the problem, we can look at the line plot of the distribution function at a convenient location in the device. The oscillations are clearly visible. The interesting observation is that these oscillations are regular; they occur every fourth grid point. Since the grid spacing in H was chosen to be such that $\Delta H = \hbar\omega_{\mathrm{opt}}/4$, this meant that the noise was somehow related to the optical-phonon energy. If in Fig. 5.9 (b) we connect every fourth point we will get a very smooth distribution function !

The cause of numerical oscillations can, therefore, be explained as follows: The Boltzmann equation with elastic acoustic phonons does not have a derivative in H. The discretized Boltzmann equation at H is only coupled to another Boltzmann equation at $H \pm n\,\Delta H = H \pm \hbar\omega_{\mathrm{opt}}$, where $n = \hbar\omega_{\mathrm{opt}}/\Delta H$. This implies that the discretized Boltzmann equations can be divided into n independent system of equations. The distribution at $H \pm \Delta H$ has no direct impact at distribution at H. This is neither numerically very promising, nor physically very sound.

Numerically, we want the the solution at neighboring points to be be coupled. If we have proper coupling then it follows that an increase in the value of $C(H + \Delta H)$ should, with other conditions remaining same, lead to an increase in $C(H)$. But in our case this would not necessarily happen, because the two grid points are not coupled. In the limiting case of $\Delta H \to 0$ it leads to absurd conclusion: If we let $\Delta H \to 0$, then the we should be able to recover $C(H \pm \Delta H) \to C(H)$; but with elastic acoustic phonons we cannot prove this in our discretized equations.

Physically too, it is not very sound. If the distribution is much larger at $H + dH$ than at H, then there has to be *some* scattering mechanism to exchange infinitesimal amount of energy to bring the distribution to equilibrium. Optical phonons cannot be such a mechanism as they possess a finite energy, $\hbar\omega_{\text{opt}}$; but *acoustic phonons in the inelastic approximation* are a viable mechanism to exchange infinitesimal amount of energy between the electrons and the lattice.

Inelastic acoustic phonons have been used in early spherical-harmonic Boltzmann work [20,32,33,47,48]. In this dissertation, the inelastic acoustic phonons were derived in a new self-adjoint form (refer to Eq. (4.20)), whose discretization couples the neighboring points in H (refer to Eq. (4.43) & Fig. 4.2 on page 65). The self-adjoint form leads to stable discretization, which provides the much-needed coupling.

The above hypothesis is corroborated by incorporating acoustic phonons in inelastic approximation in our simulation. The resulting distribution function, as shown in Fig. 5.19 on page 106, is now noise-free.

5.9 Chapter Summary

In this chapter the matrix solver for the Boltzmann equation was introduced. The four coupled equations were solved self-consistently by a Gummel-type decoupled scheme. A $\mathrm{n^+nn^+}$ device was simulated by the simulator. The simulator was also tested on a the thermal-equilibrium test case, as well its convergence characteristics were analyzed. Numerical noise in the distribution function was analyzed and it was shown that it could be eliminated by *inelastic* acoustic phonons.

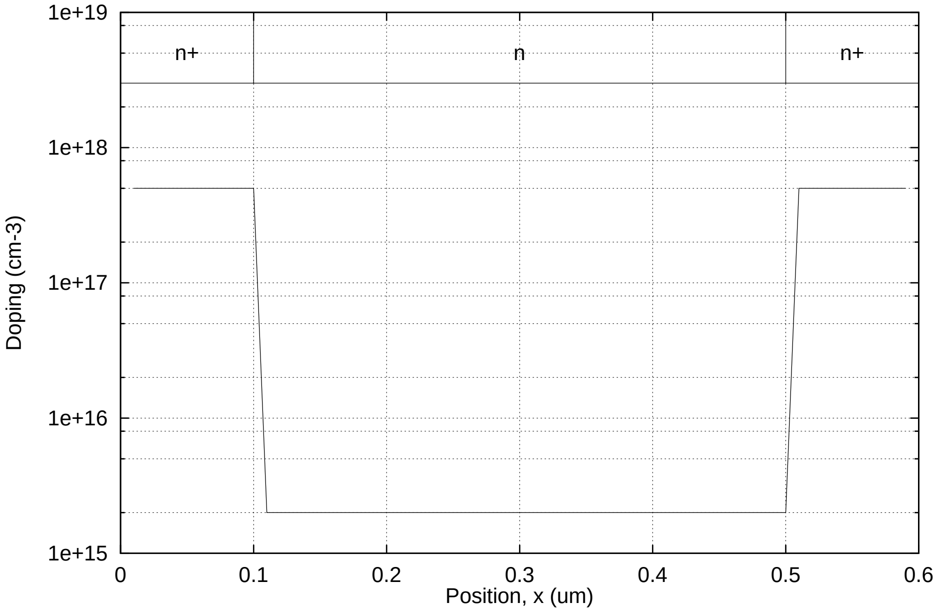


Figure 5.3: Structure of the simulated n$^+$nn$^+$ device. It has two n$^+$ regions of 0.1μm width and doping of 5×10^{17} cm^{-3} sandwiching a lightly-doped n-region of width 0.4μm and doping 2×10^{15} cm^{-3}. Ohmic contacts are at $x = 0$ and $x = 0.6\mu$m. Contact at $x = 0$ is always at ground; bias is applied to the contact at $x = 0.6\mu$m.

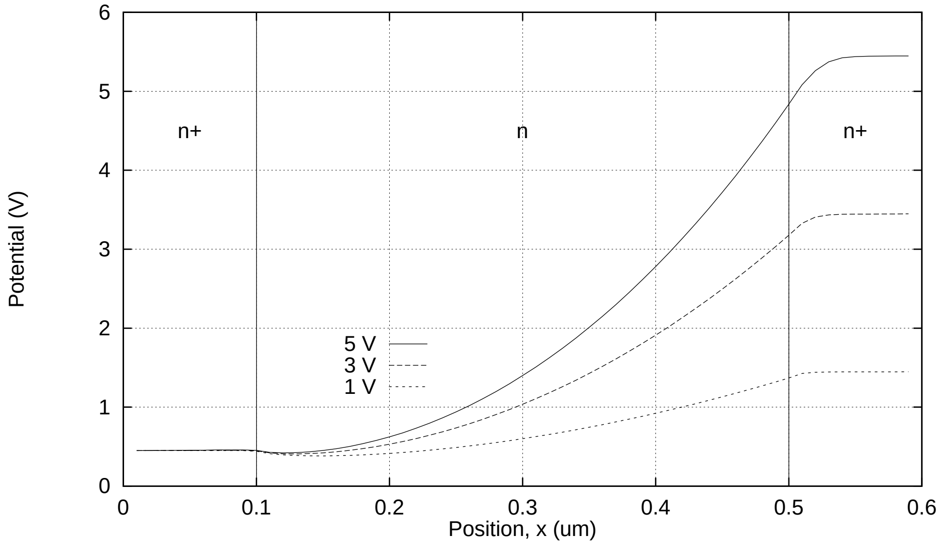


Figure 5.4: Electric potential in n$^+$nn$^+$ device as function of position for three applied biases. Most of the potential drop occurs in the lightly-doped middle region.

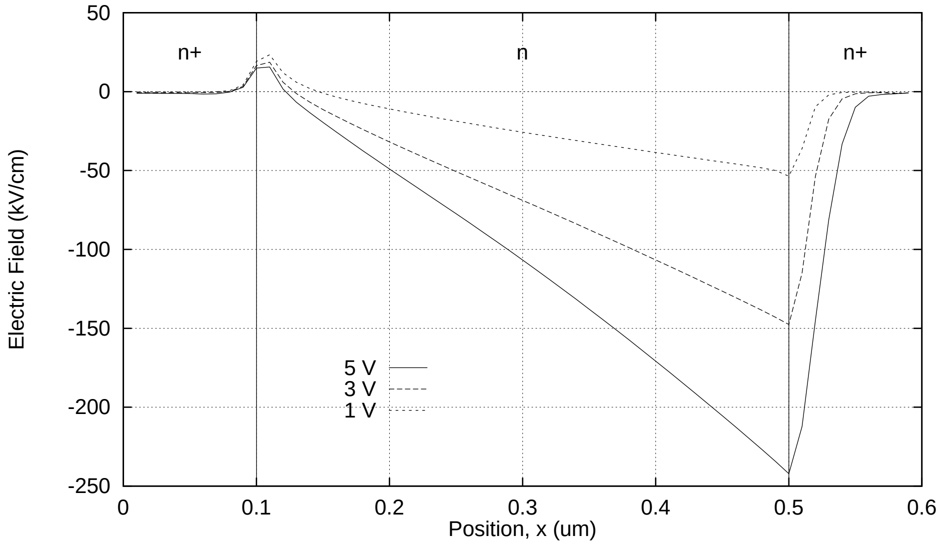


Figure 5.5: Electric-field profile within the n$^+$nn$^+$ device for three applied biases. Heavily-doped regions are field-free, and the electric field achieves a maximum near the right n-n$^+$ junction.

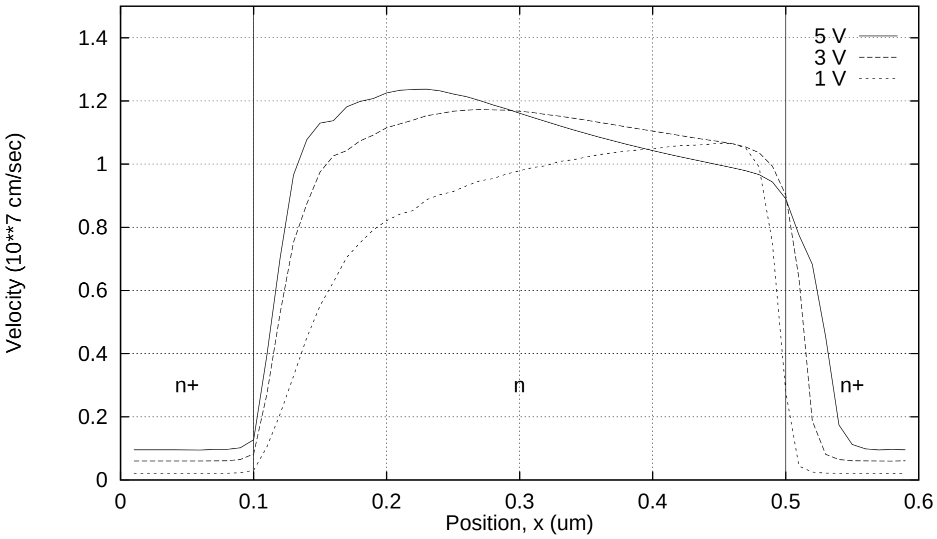


Figure 5.6: Average electron velocity in the n$^+$nn$^+$ device. Velocity overshoot is present near the "source" junction for 5 & 3V applied bias. While velocity is small in the low-field highly-doped regions, it tends to saturate at the saturation value of $\sim 10^7$ cm/s.

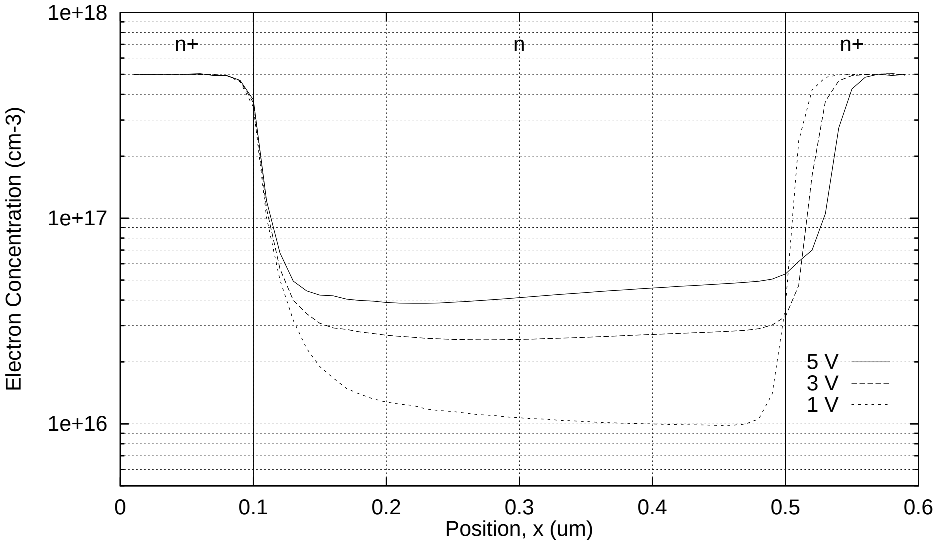


Figure 5.7: Electron concentration in the n$^+$nn$^+$ device at various applied biases.

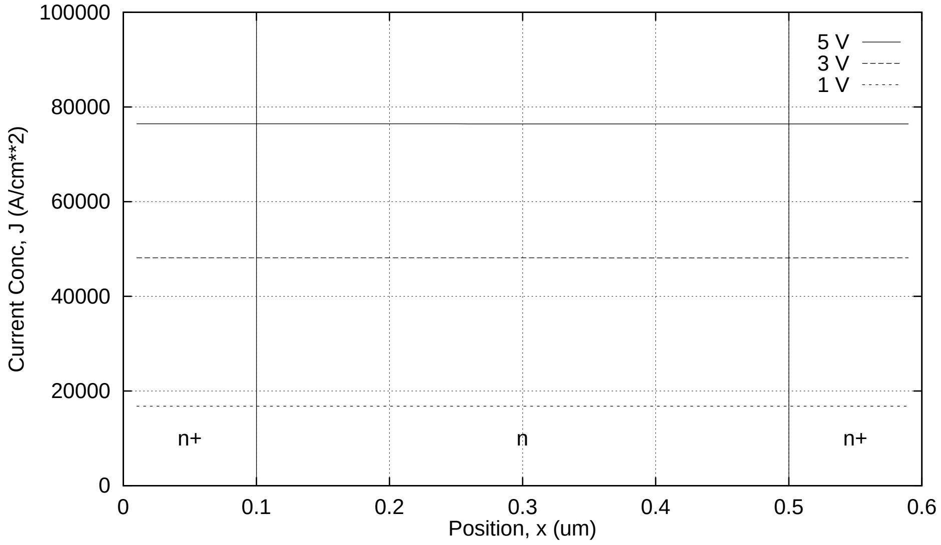


Figure 5.8: Current density as function of position at various applied biases in the n^+nn^+ device. Constant current validates the current-conserving property of control-volume formulation.

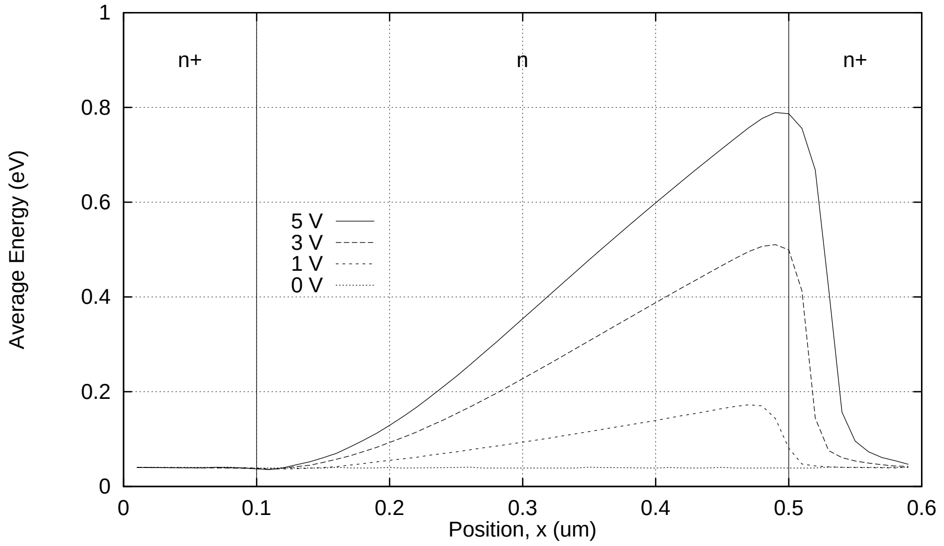


Figure 5.9: Average electron energy at various biases in the n$^+$nn$^+$ device. In the "source" and "drain" region the average energy is at the equilibrium Maxwellian value of $3/2k_BT_L$, which is 0.0388eV at room temperature of 300K.

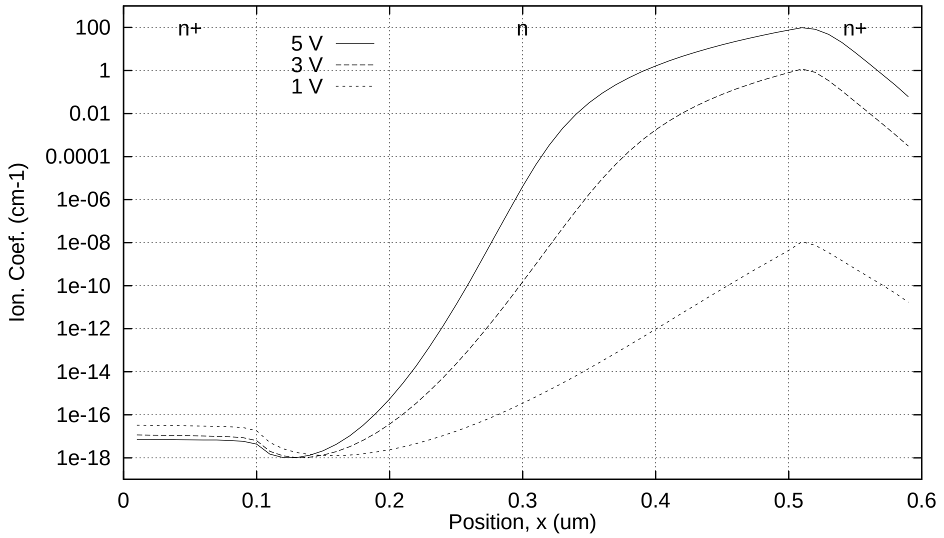


Figure 5.10: Impact ionization coefficient, α_{impact}, in the n$^+$nn$^+$ device at various biases. The peak ionization occurs beyond peak electron energy or peak electric field.

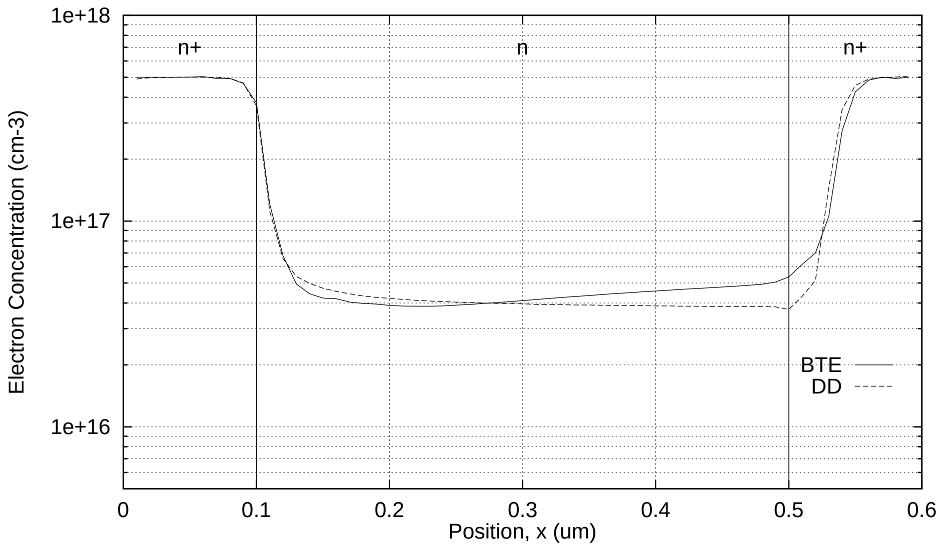


Figure 5.11: Electron concentration in a n$^+$nn$^+$ device at 5V. The curve marked "BTE" was obtained by the solution of the self-consistent Boltzmann equation. The curve marked "DD" was obtained by the solution of the drift-diffusion equations.

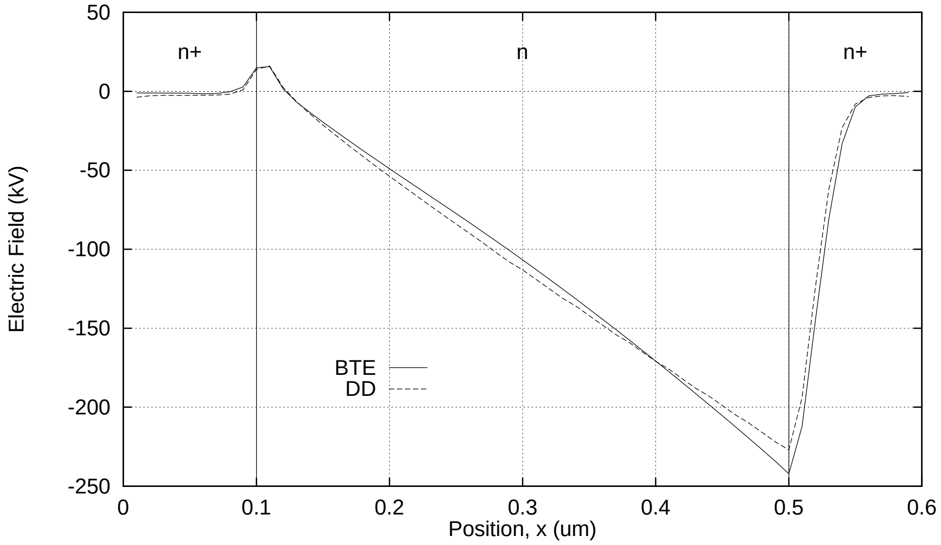
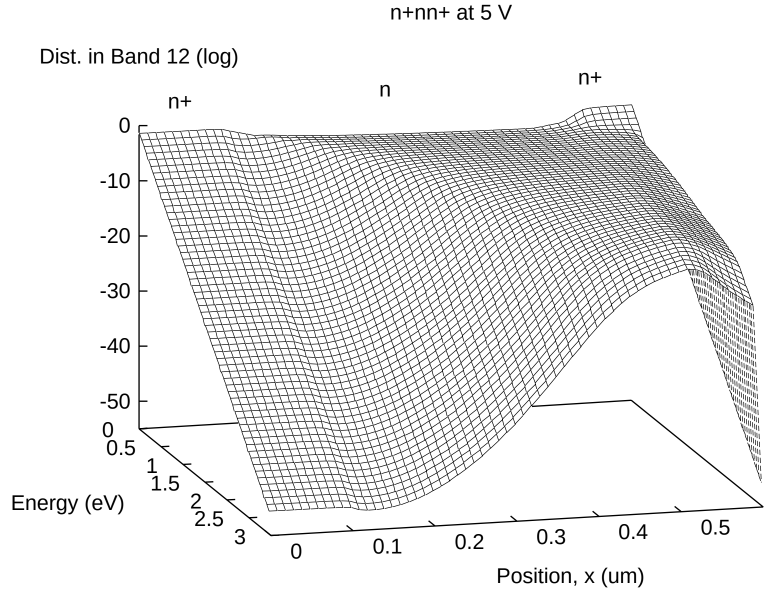


Figure 5.12: Electric field in the n$^+$nn$^+$ device at 5V. The curve marked "BTE" was obtained by the solution of the self-consistent Boltzmann equation. The curve marked "DD" was obtained by the solution of the drift-diffusion equations.



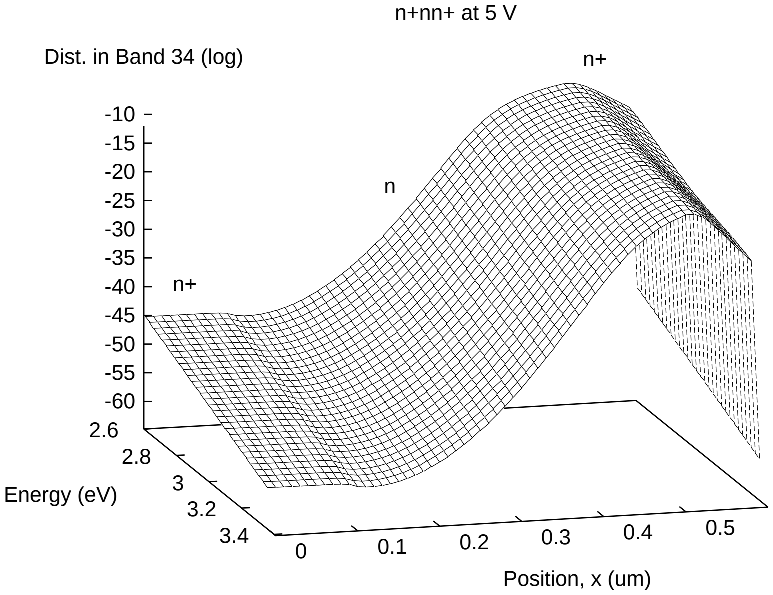


Figure 5.13: Distribution function in the n$^+$nn$^+$ device at a bias of 5V. Distribution in the lower bands, band (12), is plotted in the top, and the distribution in the upper bands, band (34), is plotted at the bottom. The electrons gain energy from the electric field and heat-up. Heated electrons scatter to bands (34) and its population increases drastically.

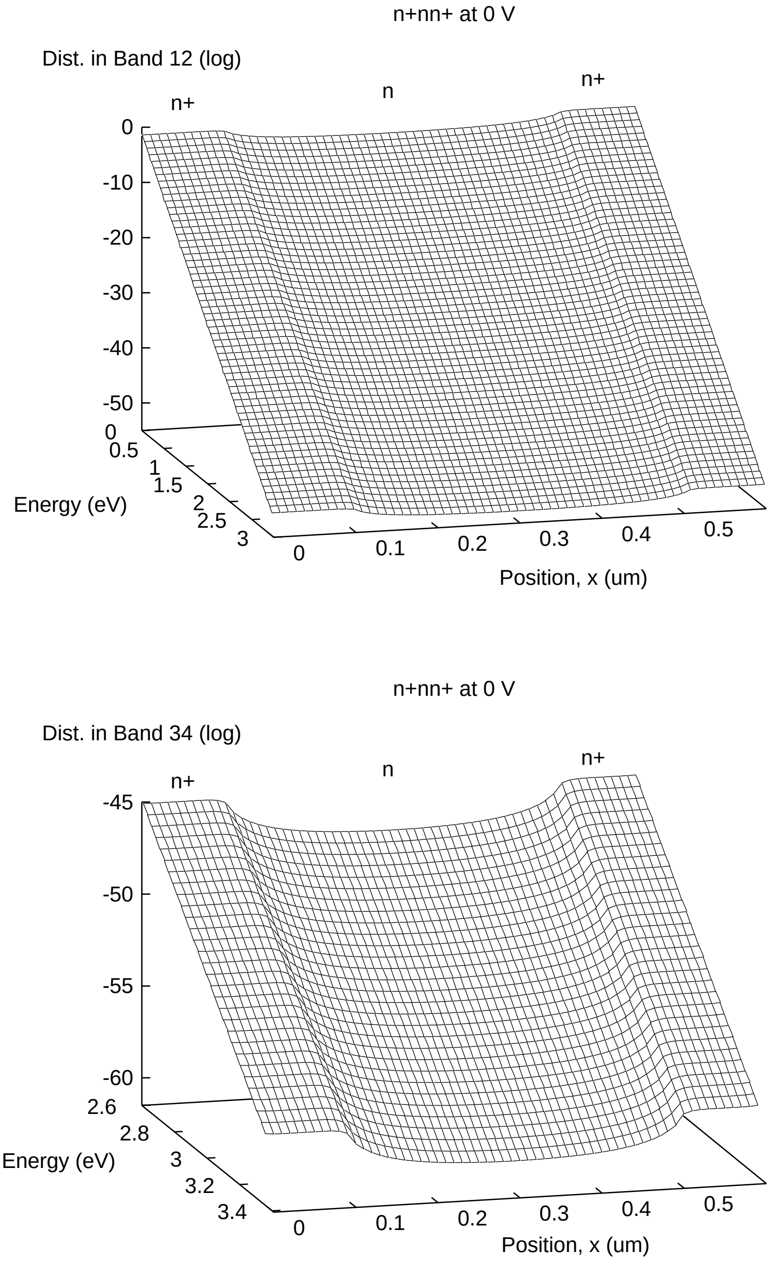


Figure 5.14: The distribution function in n$^+$nn$^+$ device at thermal equilibrium (zero applied bias). Distribution in the lower bands, band (12), is plotted in the top, and the distribution in the upper bands, band (34), is plotted at the bottom. The distribution function is Maxwellian ($\propto \exp(-\varepsilon/k_B T_L)$) in both the bands.

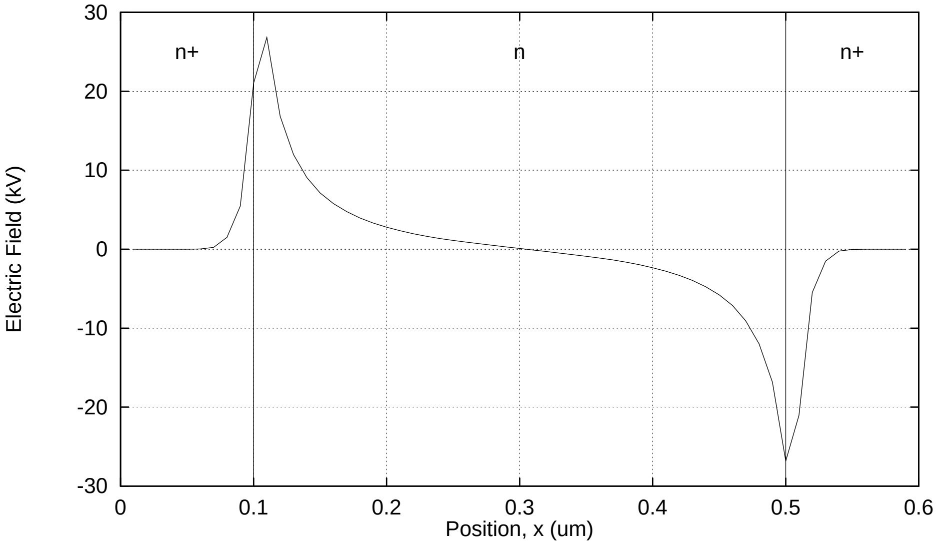


Figure 5.15: Electric field in the n^+nn^+ device at thermal equilibrium (zero applied bias). There are no currents in the device despite the presence of large electric fields.

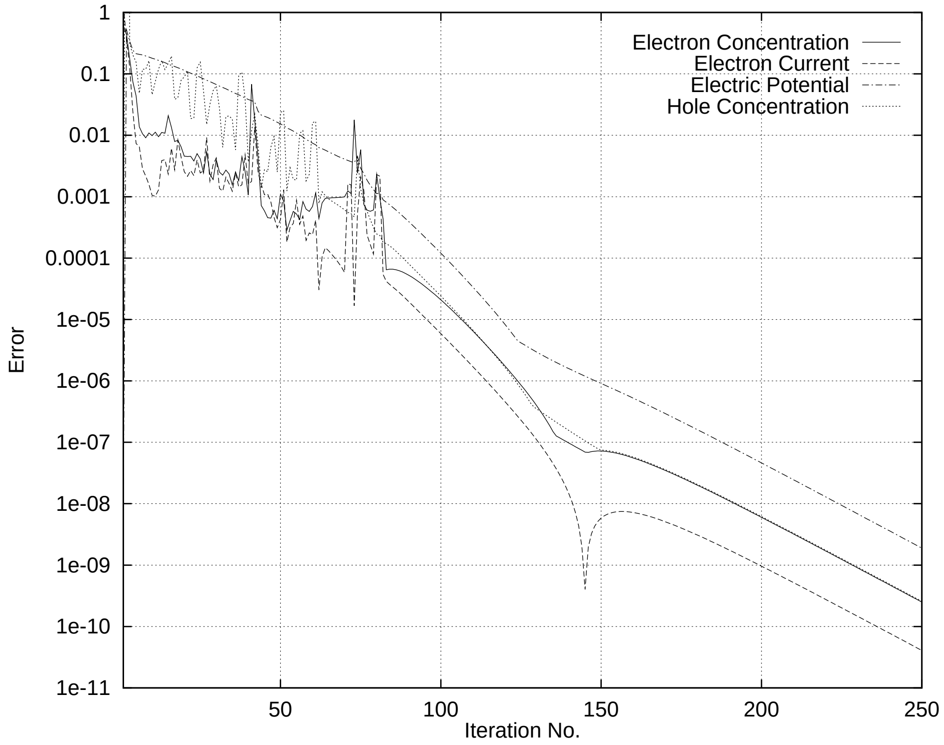


Figure 5.16: Convergence behavior of the self-consistent Boltzmann system. The error criteria for Boltzmann, Poisson, and hole-continuity equations are plotted against the outer Gummel loop. This simulations are for the n^+nn^+ device at 5V; the successive over-relaxation (SOR) factor $\omega = 1.4$. Error criteria for the Boltzmann equation is the relative change in electron concentration (marked "Electron Concentration" in the plot), and the relative change in electron current density (marked "Electron Current"). Error criterion for the Poisson equation is the change in electric potential divided by the thermal voltage (marked "Electric Potential"). Error criterion for the hole-continuity equation is the relative change in the hole concentration (marked "Hole Concentration").

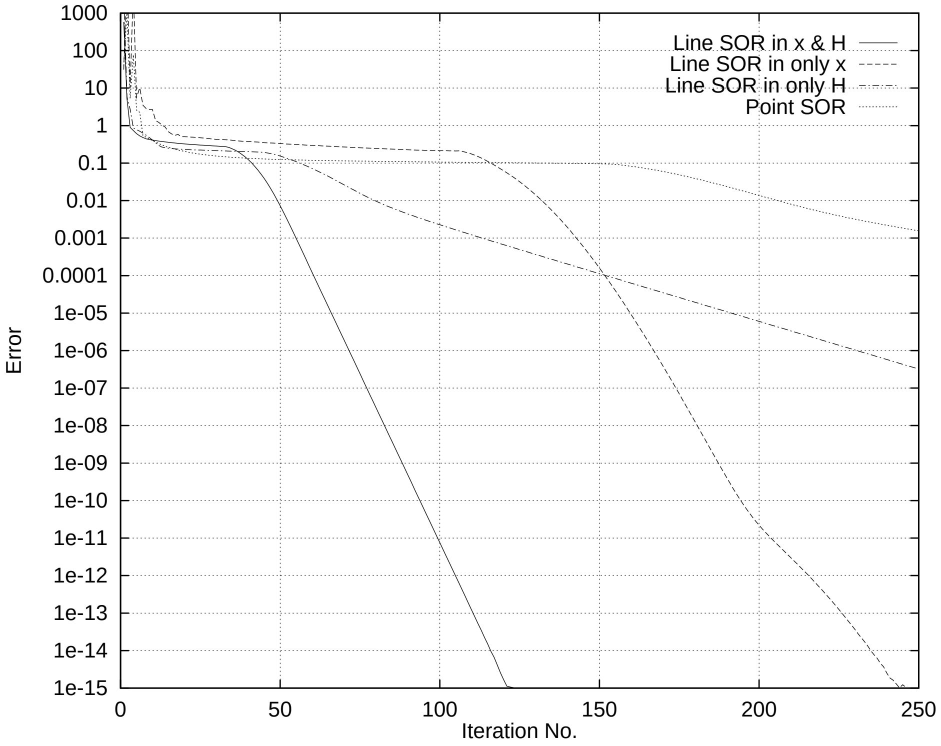
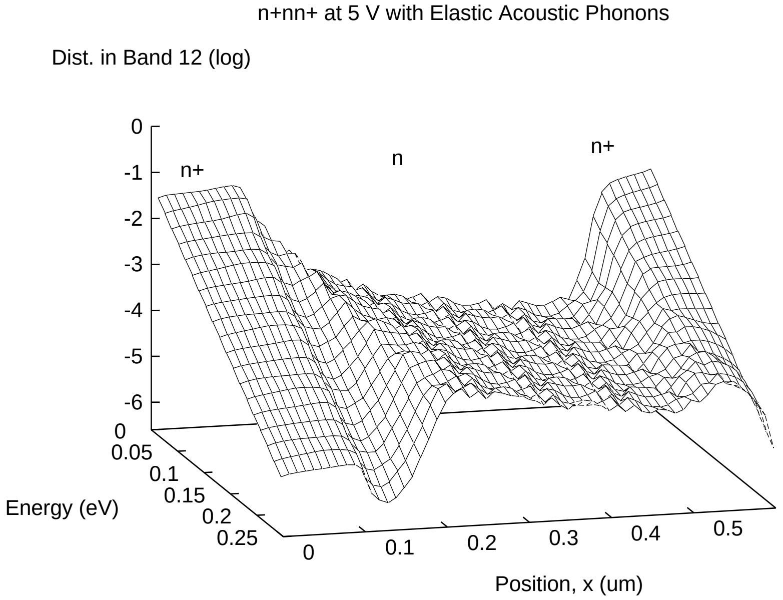
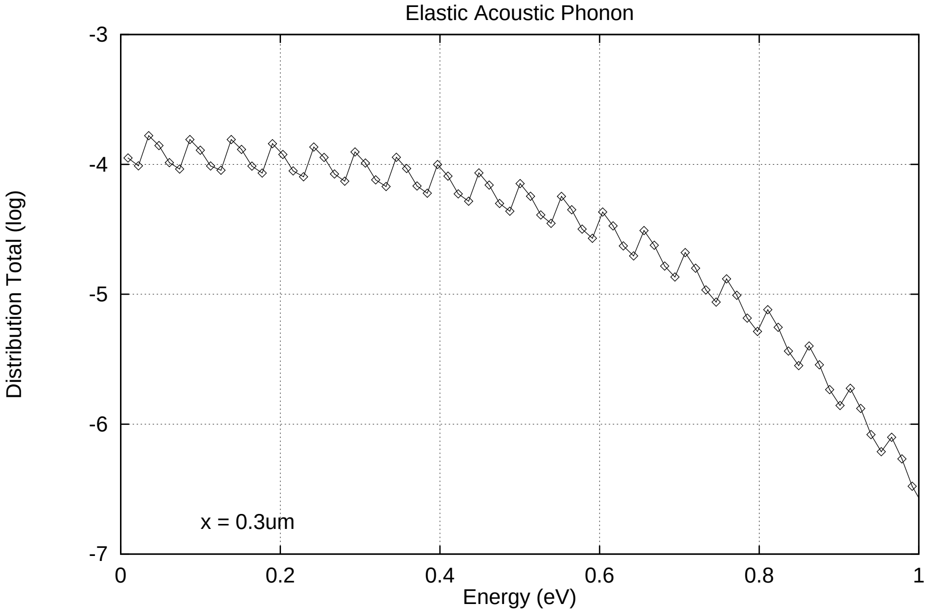


Figure 5.17: Convergence behavior of various matrix solutions of Boltzmann equation. Boltzmann equation in band (12) is solved non-self-consistently (electric potential from drift-diffusion) at a bias of 5V. Successive over-relaxation (SOR) factor $\omega = 1.2$ for all solvers. Error for the Boltzmann equation is plotted against the iteration number of the inner loop. Error in this case is the relative change in $C^{(12)}(x, H)$, and is plotted for four matrix solvers.

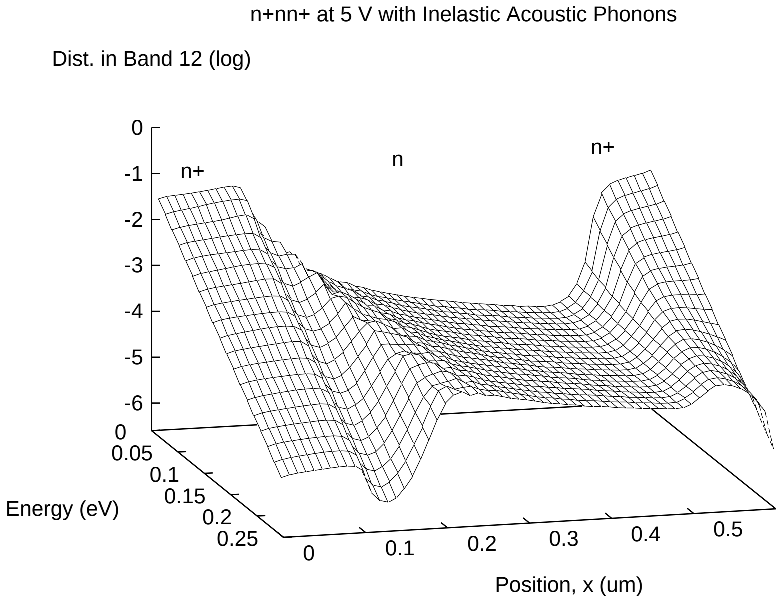


(a) Surface plot of the distribution function.



(b) Line plot of the distribution function at 0.3μm.

Figure 5.18: Distribution function obtained by considering acoustic phonons in elastic approximation. The $\mathrm{n}^{+}\mathrm{nn}^{+}$ device is at 5V bias. The grid spacing in Hamiltonian is related to the optical-phonon energy by $\Delta H = \hbar\omega_{\mathrm{opt}}/4$. The surface plot shows the *noisy* distribution. Line plot reveals further details by plotting the distribution at every grid point.



(a) Surface plot of the distribution function.

(b) Line plot of the distribution function at 0.3μm.

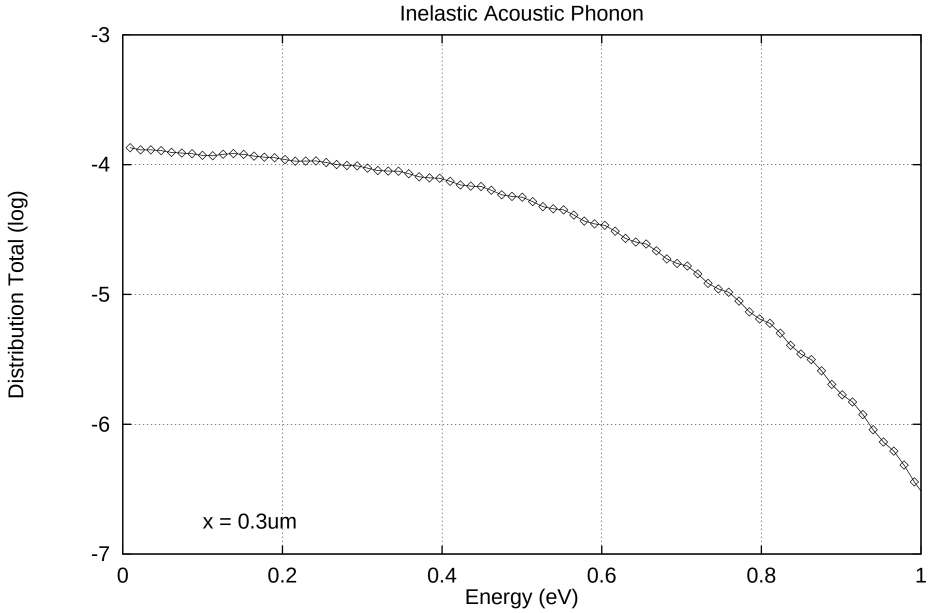


Figure 5.19: Distribution function obtained by considering acoustic phonons in inelastic approximation. The distribution function, when compared to the one with elastic approximation, is free of noise. The n$^+$nn$^+$ device is at 5V bias. The grid spacing in Hamiltonian is related to the optical-phonon energy by $\Delta H = \hbar\omega_{\text{opt}}/4$.

CHAPTER 6

SIMULATION OF BIPOLAR JUNCTION TRANSISTOR

6.1 Chapter Introduction

The numerical discretizations introduced in Chapter 4 on page 53 were applied to the simulation of a n$^+$nn$^+$ device in Chapter 5 on page 81, where the two multi-band Boltzmann equations were solved self-consistently with the Poisson and hole-continuity equations by a decoupled Gummel-type method. n$^+$nn$^+$ is a test device, often used to validating modern device simulators [109]. In this chapter we apply the same numerical techniques to simulate a npn bipolar junction transistor (BJT).

Choice of BJT is not accidental, there are five main reasons governing this choice. Firstly, BJT, after MOSFET, is one of the most common device in silicon technology. Secondly, it is essentially a one-dimensional device, making it a good candidate for our one-dimensional simulator. Thirdly, transport in a BJT, as opposed to a MOSFET and n$^+$nn$^+$ device, is essentially bipolar (two-carrier) in nature: transport is governed by both holes and electrons. Simulating hole-transport, by the hole-continuity equation, is no longer an ornamental addition—it is a necessity. Fourthly, BJT is an interesting device to simulate in its own right; there is a richness of physical detail to study. It contains an important devices like p-n junction diodes as a sub-element. And lastly, from the numerical standpoint, BJT device is a challenging device to simulate for the following reasons: very severe coupling among the equations, leading to complex convergence problems; there are very large changes in electron and hole densities over short distances; large electric-fields with rapid variations (large gradient); tendency of the electron-hole plasma to form in large part of the device (p. 259 [63]).

Bipolar transistors have been simulated by drift-diffusion and hydrodynamic or energy transport for a long time [62, 72, 86]. Simulation by solving the Boltzmann transport equation using Monte Carlo [62, 87–90, 92–94] or scattering-matrix approach [91] is relatively new. Spherical-harmonic technique has been applied to simulate one-dimensional BJT's [34,36,37,44]. Self-consistent simulations, with one conduction band, were performed in [34]; whereas, lower two bands from [57] were used in [36,37] in a non-self-consistent solution. In this chapter all the bands of the bands are incorporated in a self-consistent simulation of a one-dimensional BJT .

This chapter is organized as follows: Section 6.2 provides a brief introduction to the bipolar transistor. Simulations details specific to BJT's, like base boundary conditions, are discussed in Section 6.3 on the facing page. Results from the simulation of BJT's are presented in Section 6.4 on the next page.

6.2 Introduction to Bipolar Junction Transistor

Bipolar junction transistor (BJT) was the first useful three-terminal semiconductor device invented [78]. It is a very prominent device used in variety of high-speed high-power circuits. Its preeminence, in recent years, has been somewhat overshadowed by metal-oxide-semiconductor field-effect transistor (MOSFET). BJT has many advantages over MOSFET technology: higher speed, higher current-drive capability, higher transconductance. But it also has many disadvantages: low input impedance, higher power-dissipation, complex processing, and higher area [78]. Nevertheless, bipolar technology continues to make progress: most notably the hetero-junction BJT technology (HBT); the marriage of bipolar with complimentary MOSFET technology (BiCMOS); and the poly-silicon emitter BJT technology [78]. These developments, and high-speed circuit needs, continue to ensure the longevity of bipolar junction transistors.

Bipolar transistor is essentially a one-dimensional device—two-dimensional effects, like current crowding, are usually parasitic in nature—therefore a one-dimensional simulation is adequate. To achieve high gains and faster speeds the base width of the BJT is kept small ($<$ 100nm). This results in large electric-field and high impact-ionization in the base-collector region. Electrons in such situation routinely gain high energies; therefore, a multi-band band-structure is required. Monte Carlo simulation of BJT's can incorporate high-energy bands, but are computationally too expensive because of the presence of retarding potential barrier (emitter-base junction) in the BJT [44, 63, 64, 67, 89]. The retarding potential, however, poses little problem for the spherical-harmonic approach.

6.3 Simulation of the Bipolar Junction Transistor

The one-dimensional simulator of Chapter 5, which was used for simulating a n^+nn^+ device, is applied to a BJT. Multi-band band-structure with the transport parameters suitable for bipolar devices was employed (refer to Section 3.2 on page 44 & Section 3.4 on page 48). The Gummel's decoupled method, as shown in the flow chart of Fig. 5.2 on page 86, was used.

Bipolar transistor requires special care for the solution to converge. Firstly, a dense grid in space is required to account for rapid variations in electric field, and carrier density. In addition, it is often necessary for the drift-diffusion to slowly ramp into the desired bias.

Emitter and collector contacts have the normal Dirichlet ohmic boundary-conditions: Eq. (2.102) for the Boltzmann equation; Eq. (D.7) for the Poisson equation; and Eq. (E.7) for the hole-continuity equation. The base-contact boundary-condition, on the other hand, requires a slightly different treatment.

6.3.1 Base Boundary Condition

For a two-dimensional BJT boundary-value problem, the base contact is a simple ohmic contact, which would be treated in normal fashion. But in a one-dimensional formulation the base-contact requires special treatment. To apply the base boundary-condition we select one grid point in the quasi-neutral base and consider it be the base contact. This point could either be the center of the quasi-neutral base, or it could be the point of maximum doping [86, 99]. At this point we apply the normal ohmic boundary-condition for the Poisson equation (Eq. (D.7)) and hole-continuity equation (Eq. (E.7)). We do not, however, apply the ohmic boundary condition to the electron Boltzmann transport equation at this point—if we were to do so, the contact will "suck" all the current. The base contact, as mentioned by [81] on pp. 355–356, is not a normal ohmic contact; it is "biased" towards one carrier.

6.4 Simulation Results

6.4.1 Device Structure & Doping

The structure of an integrated-circuit two-dimensional BJT is shown in Fig. 6.1 on page 117. Since BJT is essentially a one-dimensional device—two-dimensional effects, like current crowding, are only parasitic in nature—it can be designed and optimized by a one-dimensional simulation. This one-dimensional simulation is on the *active region* of the device. The active region, as shown in Fig. 6.1, is in the vertical dimensional; it starts from the base, goes through the emitter ending in the collector. This is the path that current takes; doping in this direction critically determines device performance.

Two BJT structures have been simulated in this work: n-type and p-type collector design. The piece-wise linear doping profiles for the two BJT is shown in Fig. 6.2 on page 117. This doping profile was used in Monte Carlo [87] as well as spherical-harmonic BJT simulations [36, 37].

The n-type BJT is the conventional collector design; maximum electric field occurs near the base. The p-type represent the other extreme; maximum electric field is located near the n^+ collector [87]. The lightly doped region of the collector, for both designs, is depleted; base width is 0.05μm.

6.4.2 Grid & Simulation Time

Grid is uniform in space and energy. There are about 200 points in space; 440 points in energy, out of which 349 were in band (12), and 91 were in band (34). Grid spacing in Hamiltonian is $\Delta H = \hbar\omega_{\mathrm{opt}}/6$. Thus, there for totally around 88,000 points for both the Boltzmann equations. The simulations were run on a 200-MHz Pentium Pro personal computer (PC) with 64 MB of RAM (Random-Access Memory) running a Linux RedHat 4.0 operating system using the simulator described for the $\mathrm{n}^+\mathrm{nn}^+$ device in Section 5.4 on page 85. The complete self-consistent solution, as shown in the flow chart of Fig. 5.2 on page 86, is obtained in around 20–30 minutes within 100 Gummel iterations.

Same bias was used in almost all the BJT simulations: Base-emitter voltage, V_{BE}, set to 0.95V; and collector-emitter voltage, V_{CE}, set to 3.95V. This bias was selected so the resulting current density was approximately $10^5\mathrm{A/cm}^2$, which meant that the electron density in the lightly-doped collector is of the same order as the doping concentration. This ensures nearly low-level injection conditions, while still keeping the biases high.

6.4.3 Electric Potential & Field

The self-consistent potential, $\phi(x)$, and electric field, $\mathbf{E}(x)$ are shown in Fig. 6.3 on page 118 for the two designs. The conventional n-type BJT has the maximum electric field at the base-collector junction ($x \sim 0.1\mu$m). For the p-type BJT the electric field is inverted, the maximum occurs at the collector-sub-collector junction ($x \sim 0.2\mu$m) [36, 37, 87]

6.4.4 Distribution Function

This multi-band energy distribution function is shown in Fig. 6.4 on page 119 for the n-type design, and Fig. 6.5 on page 120 for the p-type design. A brief explanation follows.

Electric motion in the BJT's is from left-to-right. In the emitter and neutral base, electric field is close to zero, therefore the distribution function is Maxwellian ($\propto \exp(-\varepsilon/k_B T_L)$). Since the electron density is lower in the base, the distribution is lower too. Near the base-collector space charge region there is large electric field: electron accelerate and gain energy from the electric field. This is seen in the shape of the distribution function, it is loosing its Maxwellian shape as electrons are accelerating to higher energy. Electrons from band (12) scatter to band (34) by inter-band optical scattering; the population of the band (34), therefore, increases by many orders of magnitude. These *hot electrons* are important, and one needs accurate multi-band band-structure to model them. These hot electrons contribute to impact ionization, which in turn determines the breakdown voltage of the BJT; this is an important hot-electron effect in BJT's. The sub-collector region is field-free; the electrons, then, loose energy to the phonons, and the distribution tends to relax to a Maxwellian. Electrons scatter from the upper bands back to the lower bands, distribution in bands (34) decreases. At the collector ohmic contact, a Dirichlet Maxwellian boundary-condition has been enforced; this causes a slight discontinuity in the distribution near the contact as the distribution has not had enough time to fully relax to a Maxwellian. We can also notice that in the sub-collector region, based on the slope of the distribution functions, there seem to be two populations in existence: one is the background sea of cold electrons, and the other is the injected hot electrons [110].

The distribution function provides complete information about the device. The average quantities, calculated from the distribution, will now be discussed.

6.4.5 Electron Velocity & Concentration

Electron velocity and concentration are plotted in Fig. 6.6 on page 121. Electron density is high in the field-free emitter region, therefore the electron velocity is negligible. In the quasi-neutral base there is no electric field, but there is a gradient of electron density, which results in a small velocity.

Electrons at the base-collector junction experience a sudden increase in electric-field. This large gradient of electric-field causes *velocity overshoot* in electrons, as explained on page 88 in Section 5.5.

There is slight, but perceptible, difference in the n-type and velocity overshoot: it is higher for the n-type design. This is a satisfying result because electric field, as shown in Fig. 6.3, at the base-collector junction is higher for the n-type design. In addition, since the electric field in p-type BJT increases with depth, electrons maintain their saturated velocity in p-type design for a more depth than n-type design.

6.4.6 Current Density

The discretization by control-volume approach was undertaken with the intention of retaining the current-conserving nature of spherical-harmonic the Boltzmann equation. This is validated in Fig. 6.7 on page 122, where we see that current density is constant.

6.4.7 Hole Concentration

As part of the self-consistent solution we solved the hole-continuity equation. Hole transport is very important in a bipolar device, it forms the majority carrier in the base as well as contributes to device operation. Hole concentration is plotted in Fig. 6.8 on page 122. We see that in the neutral base, where holes are the majority carrier, hole concentration is the highest. A sharp decrease in hole concentration is seen near the emitter and collector ohmic contact, this is because of the Dirichlet boundary-condition on hole concentration.

6.4.8 Average Energy

Average electron energy in the n-type and p-type BJT's is plotted in Fig. 6.9 on page 123. The p-type BJT has a larger average energy, this is due to an electric-field profile which is slowly varying and has a large magnitude, thus allowing the electrons to heat up. Also, the average-energy profile in the p-type collector follows the electric field profile closely, this is because of smaller gradient of of the electric field. In fact it actually peaks 0.012μm before the electric field, this is due to the diffusion of cold electrons from the heavily-doped sub-collector. For the n-type design, average energy peaks 0.057μm after the electric field maximum. This delay is known as the *dead-space effect* [94], or sometimes as the *dark-space effect* [69]. This non-local behavior was also observed in Monte Carlo simulation of the two BJT's [87].

6.4.9 Impact Ionization

Fig. 6.10(a) on page 124 plots the impact-ionization coefficient, α_{impact} for the n-type and p-type BJT. The two profiles are very different. Owing to the larger average energy achieved, the impact-ionization coefficient of p-type design is much larger than n-type design. Due to non-local, or non-equilibrium, nature of transport the impact ionization peak does not necessarily coincide with the peak of either the electric field or average energy [87, 94]. The delay of impact ionization is also known as the *dead-space effect* [94], or as the *dark-space effect* [69].

In both designs the impact-ionization coefficient peak does not coincide with the respective average-energy peak: For the p-type BJT it occurs 0.014μm deeper; and for the n-type BJT it occurs 0.015μm deeper. This difference can be attributed to the fact that average energy is not necessarily a perfect representation of the hot-electron effects like impact ionization. Average energy is an averaging of the distribution over all energies, while impact ionization depends only on the high-energy tail. A sea of cold electrons, like the heavily-doped sub-collector, can result in lower average energies [87].

Using electric field as a measure of impact ionization is even more inaccurate. For the n-type collector design, maximum impact ionization occurs 0.072μm after the maximum electric field. The mismatch is not so severe for the p-type design; the two maxima are virtually coincident. The two BJT behave differently because of the n-type design has larger gradient of the electric field at the base-collector junction. An impact ionization model based on the local electric field would tend to be erroneous in such situations. Fig. 6.10(b) on page 124 shows the impact-ionization coefficient computed from the local electric field in a Chynoweth model [72]. The n-type impact-ionization coefficient actually peaks at the wrong side of the depletion region!

6.4.10 Effect of Including Impact Ionization

Fig. 6.11 on page 125 shows the effect of including impact ionization in the simulator. Total distribution function is plotted at two locations in the n-type BJT at two locations. The total distribution function is the weighted sum of distribution in band (12) and band (34), where weighing is done by density of states in the two bands.

Fig. 6.11(a) plots the total distribution at a point in the high-electric-field collector region. At this point electrons are very hot, and impact ionization is high. There is a loss of electrons form the high-energy tail due to impact ionization.

Fig. 6.11(b) plots the distribution at a point in the field-free sub-collector region. Presence of impact ionization drives the distribution towards equilibrium (Maxwellian) faster.

6.4.11 Effect of Including Ionized Impurity

Ionized impurity scattering is more pronounced for low-energy carriers (refer to Fig. 3.4 on page 51 and [62]). It affects, therefore, low-energy transport properties. Velocity overshoot is a low-energy phenomena: when low-energy electrons suddenly experience a high gradient of electric field, they overshoot their saturation velocity [107]. The low-energy velocity is higher because the low-energy scattering is lower (or low-energy mobility is higher) [107]. (Refer also to page 88 in Section 5.5.2).

Scattering rate at low-energies is increased by including ionized impurity scattering. Fig. 6.12 on page 126 plots the average electron velocity in the n-type BJT with and without the impurity scattering. When we include this scattering, three things are obvious: (a) Velocity overshoot is less pronounced; (b) velocity in neutral base, which is due to low-energy near-Maxwellian electrons, is much lower; and (c) saturation velocity, which is a high-energy balance of electric field with optical phonons, is not affected.

6.4.12 Comparison with Drift-Diffusion

Fig. 6.13 on page 126 compares the Boltzmann-equation solution with the drift-equation (DD) solution. The electron density from the two models differs in the collector region, this is possibly because drift-diffusion neglects the thermal component in the electron current. In addition, near the base-collector junction ($\sim 0.1\mu$m) electron density is lower in the Boltzmann model; this is dip occurs due to velocity overshoot in Boltzmann model, which is absent in the drift-diffusion model.

6.4.13 Ballistic Transport

Ballistic transport occurs in small semiconductor devices where electrons may travel short distances without scattering [108]. An electron traveling ballistically for a distance Δx in an electric field E gains kinetic energy $\Delta\varepsilon$ which, from the usual Newton's laws, is given by

$$\Delta\varepsilon = -q\,E\,\Delta x = q\,\Delta\phi \tag{6.1}$$

where $\Delta\phi$ is the change of potential over this distance. If all electrons undergo ballistic transport, then the distribution just shift in energy; therefore, the distribution at $x_0 + \Delta x$ is

$$f_0^0(x_0 + \Delta x, \varepsilon + \Delta\varepsilon) = f_0^0(x_0 + \Delta x, \varepsilon + q\,\Delta\phi) = f_0^0(x_0, \varepsilon) \tag{6.2}$$

This signature of ballistic transport—distribution displaced in energy—is present in the high-field base-collector space-charge region of the BJT, as shown in Fig. 6.14 on page 127. The figure plots the total distribution at three locations: $x = 0.11, 0.12$, and 0.13μm; where the potential is $\phi(x) = 0.95$, 1.38, and 1.73V respectively. The distribution shows two distinct slopes: a tail region in equilibrium with the lattice ($\propto\ \exp(-\varepsilon/k_BT_L)$); a low-energy region at a very high effective temperature (almost flat in energy). In addition, there is a distinct breakpoint between the two regions. This breakpoint gives a marker to measure the shift of the distribution in energy. We see that the distribution are shifted in energy by an amount $\Delta\varepsilon = q\,\Delta\phi$ from one another, which proves the existence of ballistic transport.

We examine ballistic transport further. Since the Hamiltonian $H = \varepsilon - q\,\phi$ the change in Hamiltonian, from Eq. (6.1), during ballistic transport is $\Delta H = \Delta\varepsilon - q\,\Delta\phi = 0$; which implies that a electron in ballistic transport conserves its Hamiltonian. This aspect of ballistic transport is illustrated in Fig. 6.15 on page 128. Plotted in this figure are the contours of electron population $f_0^0(x, \varepsilon)g(\varepsilon)$, which is proportional to distribution times the density of states. The Hamiltonian H is the y-axis. Bending of the conduction band in the base-collector space-charge region indicates a region of high electric field region. Electrons entering the region between the two arrows (region labeled "ballistic transport") experience a sudden electric field. For a short distance the electrons that enter this region are accelerated by the field ballistically; without scattering. This ballistic transport is revealed by the contours of electron density (between the arrows) are flat in H even though the conduction band is bent; $\Delta H = 0$ for the electrons. This ballistic transport ends, and electrons scatter and loose energy; this is indicated by the contours bending downward beyond the arrows.

6.4.14 Graded Doping in Base

In a uniformly-doped base there are no electric fields, so electrons motion in neutral base is diffusion dominated. A non-uniformly doping produces a built-in electric field in the base, which helps accelerate electrons [80]. Fig. 6.16 on page 129 illustrates the effect of base doping profiles. Fig. 6.16(a) shows the three doping profiles with varying gradient in the base. Doping gradient causes a built-in electric field ($\propto\ (dp/dx)/p = d\ln p/dx$), which accelerates electrons. Fig. 6.16(b) plots the corresponding average velocities. Velocity for electrons in the neutral base is higher for larger doping gradient. This improves cut-off frequency since the base-transit time is reduced.

It can also be seen in Fig. 6.16 velocity overshoot is less pronounced if the electric field is less sudden (less gradient of electric field).

6.4.15 Thermal Equilibrium

Thermal equilibrium, as discussed in Section 5.6 on page 89, is a very important test case. It is proved in Appendix B on page 177 that at thermal equilibrium the distribution function in a device is Maxwellian. Fig. 6.17 on page 130 plots the distribution function for both the bands of our band-structure. We can see that the distribution function in both bands is a straight line on the log scale: It is Maxwellian at lattice temperature (it is $\propto\ \exp(-\varepsilon/k_BT_L)$).

The Maxwellian nature of the of the distribution is further proved by the plot of the average energy in the device in Fig. 6.18(a) on page 131. The average energy of a Maxwellian distribution is $3/2\,k_BT_L$ (p. 106 of [74]); which at room temperature of 300K is approximately 0.0388eV. In Fig. 6.18(a) the average energy is virtually constant at this thermal-equilibrium value, thus further proving that the distribution function is Maxwellian.

It is interesting to note that equilibrium conditions hold despite the presence of large electric fields in the device. Fig. 6.18(b) shows the built-in electric fields at zero bias conditions. There is large electric field at the emitter-base junction ($>$ 800kV/cm), and base-collector junction. These fields cause no current in the device: drift and diffusion currents cancel out exactly. Electron velocity, not plotted, was found to be virtually zero throughout the device. This, as pointed out in Section 5.6 on page 89, can be interpreted as *velocity undershoot* or *velocity damping*: electron velocity is much smaller than what is to be expected from the local electric field.

6.5 Chapter Summary

Numerical techniques developed in the previous chapter were applied to simulate a one-dimensional short-base bipolar junction transistor (BJT). Simulation were performed by a self-consistent solution of four equations: Two Boltzmann equations for all bands of the multi-band band-structure; Poisson equation; and hole-continuity equation. Two BJT structures were analyzed. Numerous physical phenomena—like ballistic transport, velocity overshoot, non-local non-equilibrium transport, impact ionization, ionized impurity, graded doping—were analyzed. The thermal equilibrium test was studied for the BJT.

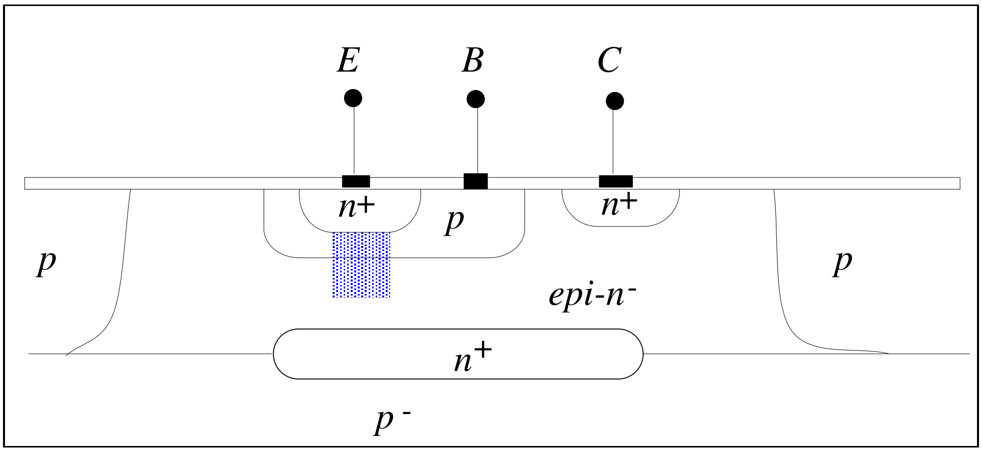


Figure 6.1: Two-dimensional structure of a planar n-p-n integrated-circuit bipolar junction transistor. For the sake of clarity, the vertical dimension is much more magnified than the horizontal dimension. Emitter ("E"), base ("B"), and collector ("C") contact are shown. Active region of the device is shown shaded. the p-type dopings on the sides are for junction isolation. The n^+ buried layer reduces the series resistance. After [80].

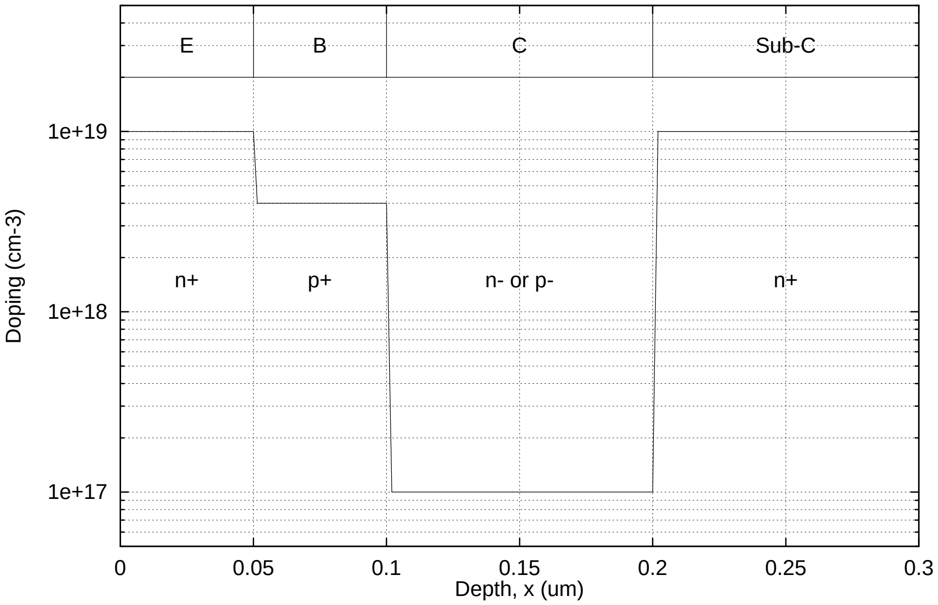
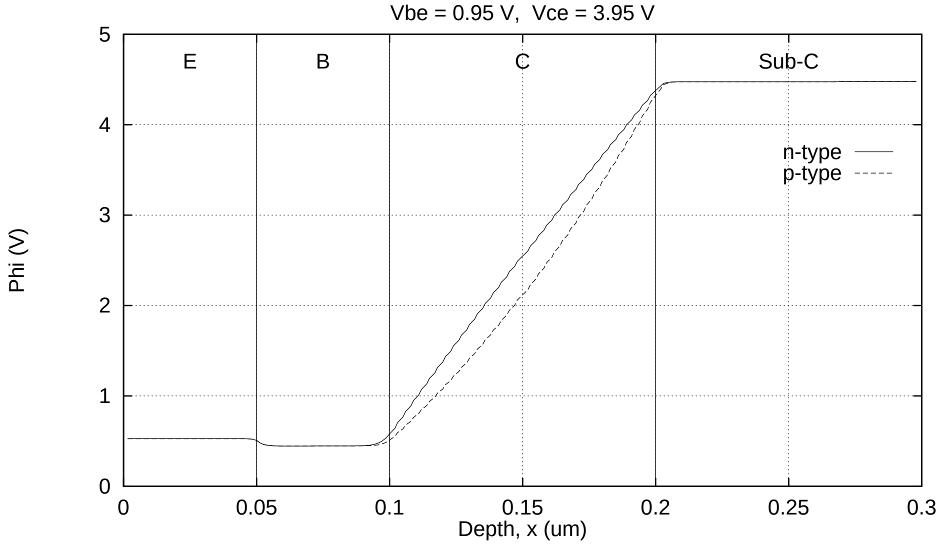


Figure 6.2: The doping profiles for the n-type and p-type BJT's. The various parts of the bipolar transistor have been marked: emitter ("E"), base ("B"), collector ("C"), and sub-collector ("sub-C"). The collector is doped differently for the n-type and p-type design.



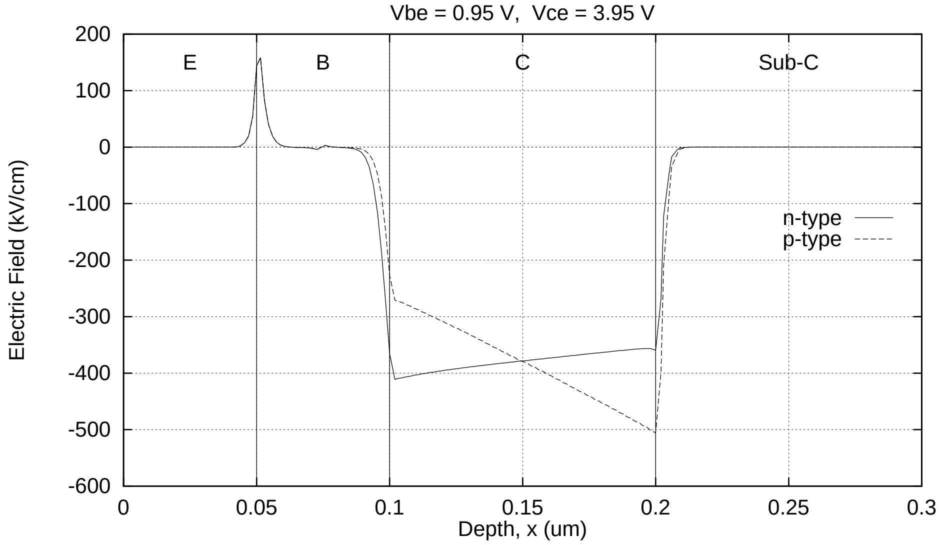


Figure 6.3: Electric potential, $\phi(x)$, and electric field, $\mathbf{E}(x)$, for the n-type and p-type BJT's. Bias is V_{BE} = 0.95V, and V_{CE} = 3.95V. This gives a V_{CB} =3.0V. The electric field is qualitatively different for the two BJT's, achieving its peak at different locations in the device.

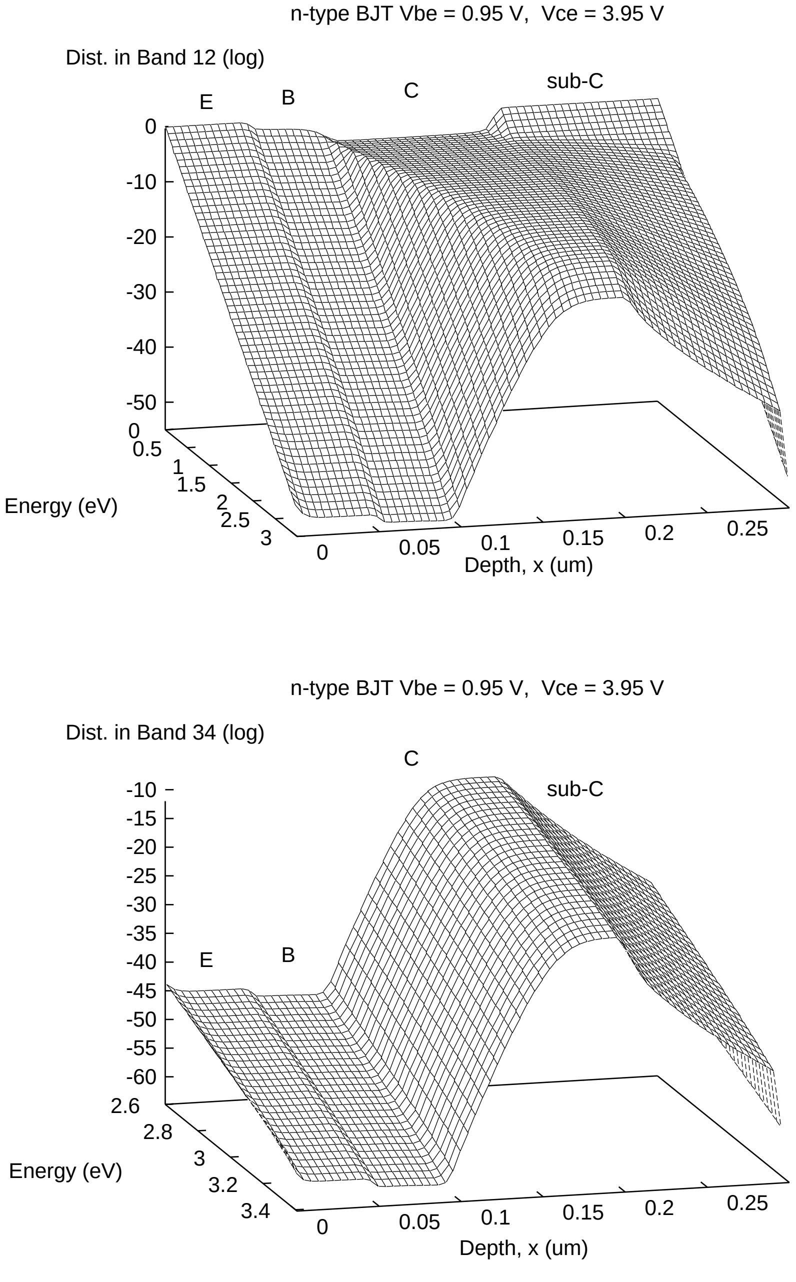


Figure 6.4: Electron energy distribution function in the n-type BJT in the bands (12) and (34). Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

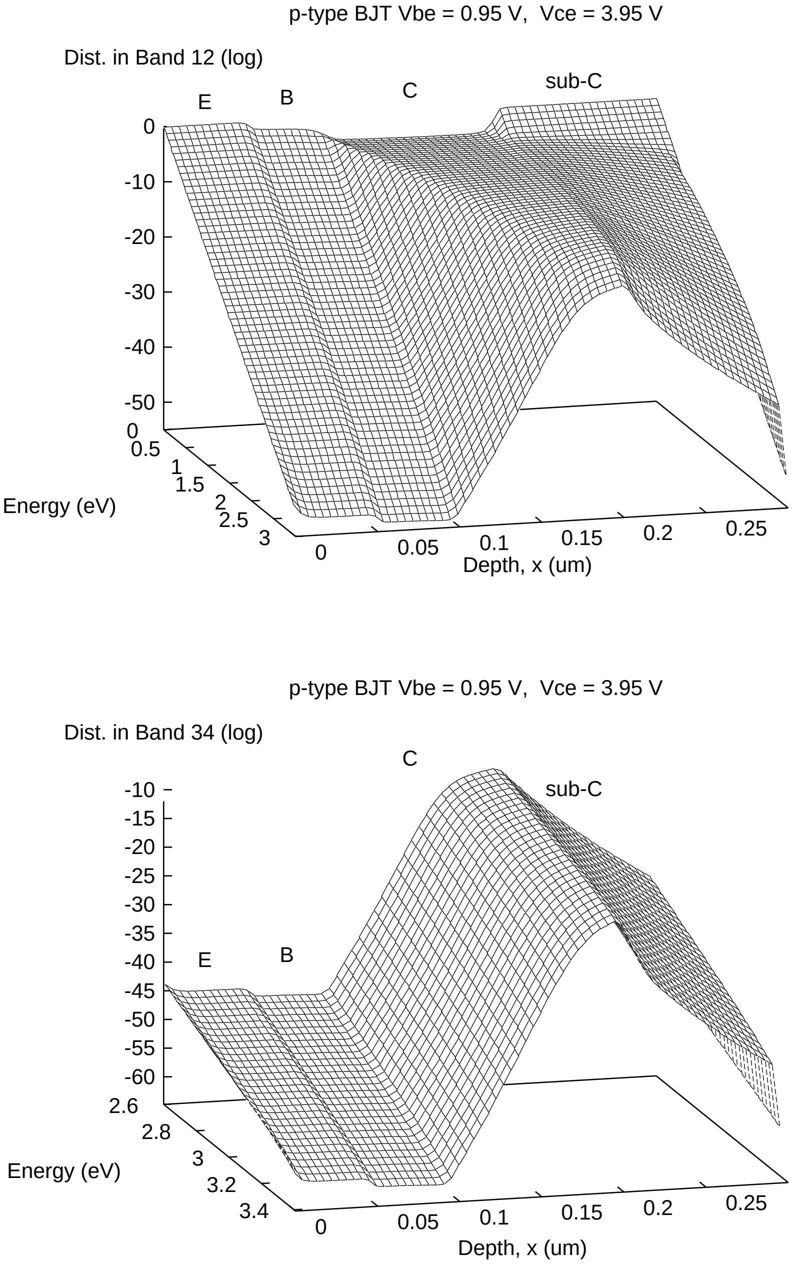


Figure 6.5: Electron energy distribution function in the p-type BJT in the bands (12) and (34). Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

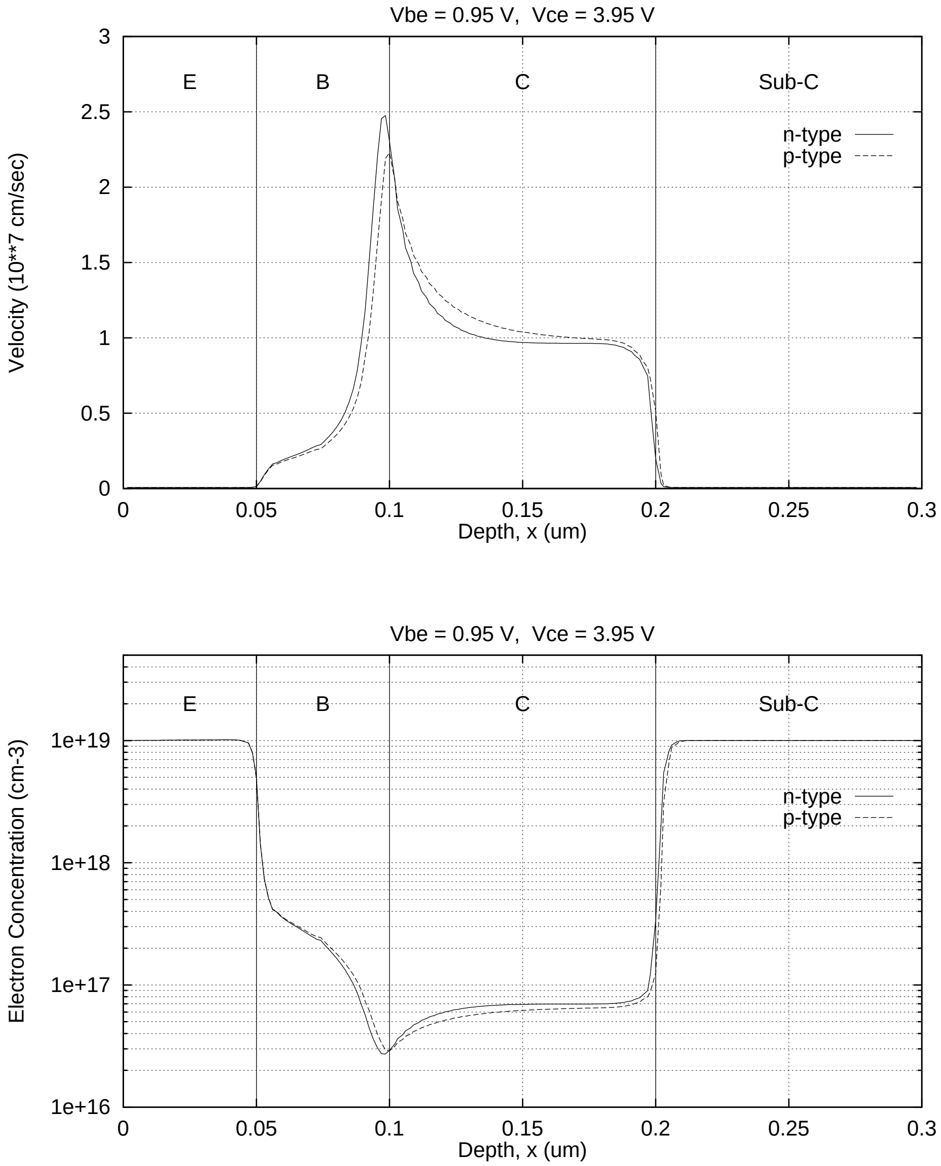


Figure 6.6: Profile of the electron velocity and electron concentration in the n-type and p-type BJT design. Bias is $V_{BE} = 0.95\text{V}$, and $V_{CE} = 3.95\text{V}$

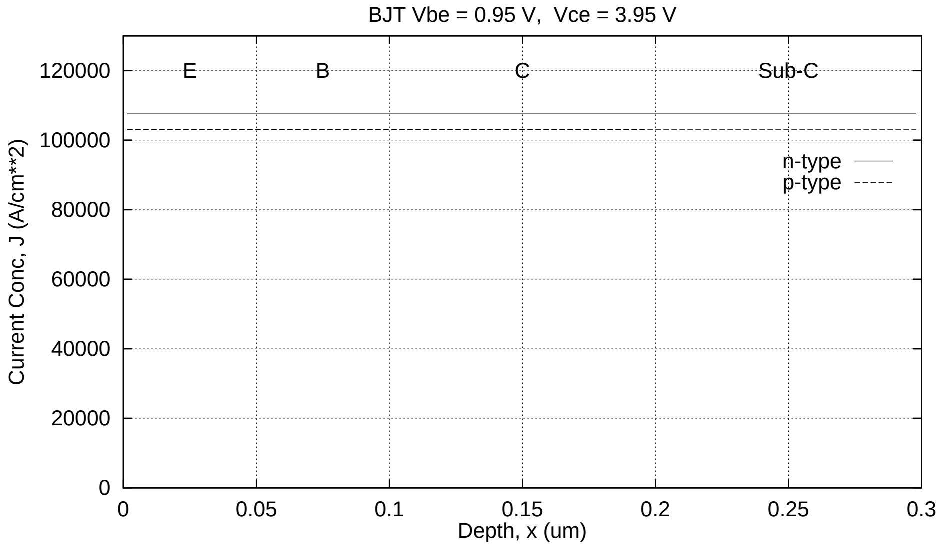


Figure 6.7: Electron current density in the n-type and p-type BJT's. The constant current demonstrates the current-conserving property of the control-volume discretization. Bias for both BJT's is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V

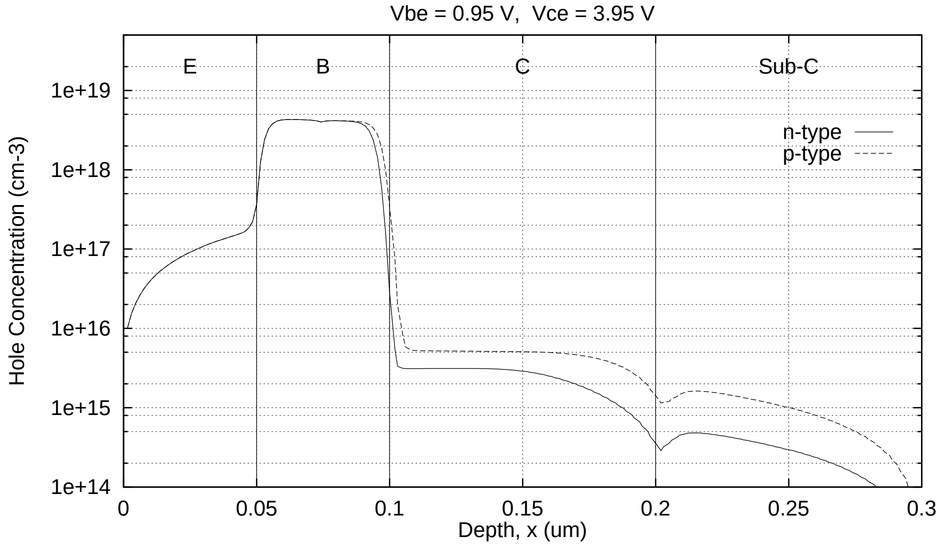


Figure 6.8: The profile of hole concentration in the n-type and p-type BJT's. The hole concentration is obtained by solving the hole-continuity equation self consistently with the Boltzmann and Poisson equations. Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

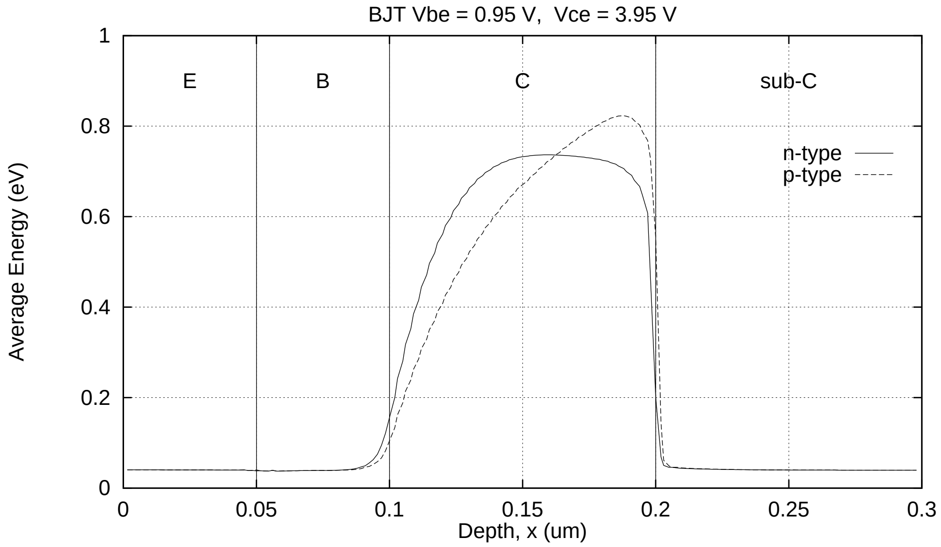
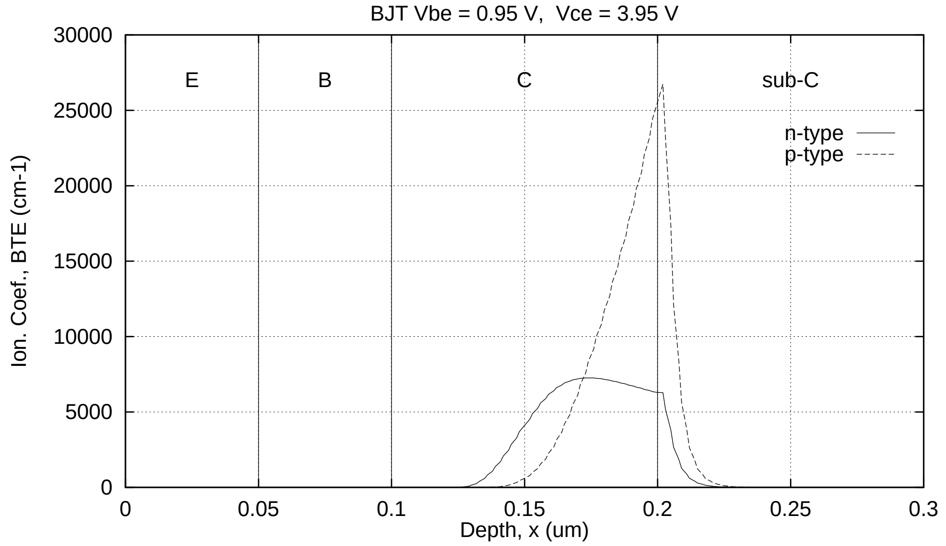
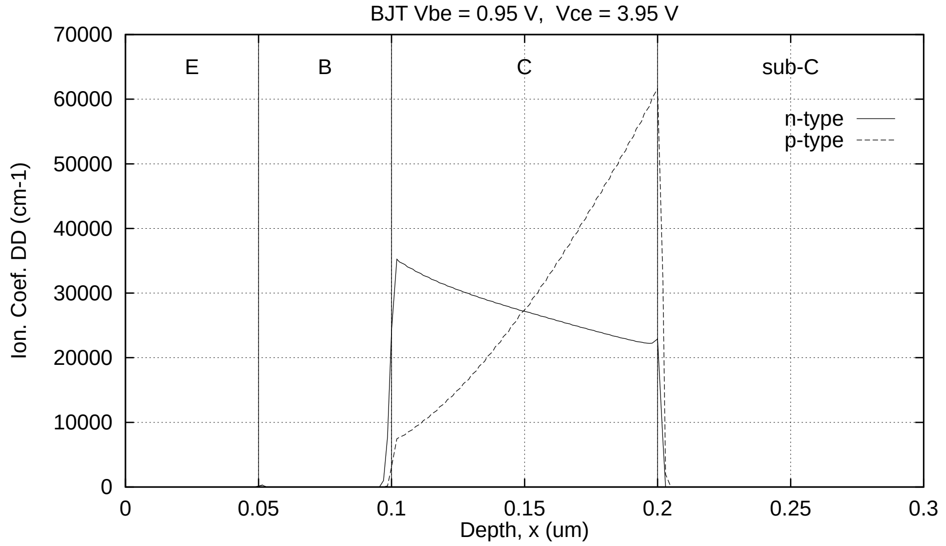


Figure 6.9: Average electron energy profile in the n-type and p-type BJT's. Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

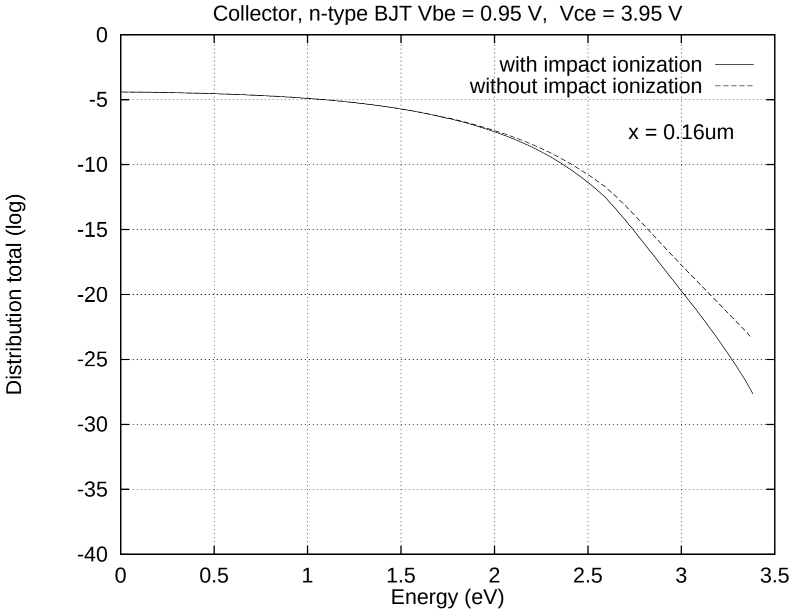


(a) Computed by using the distribution function after solving the BTE

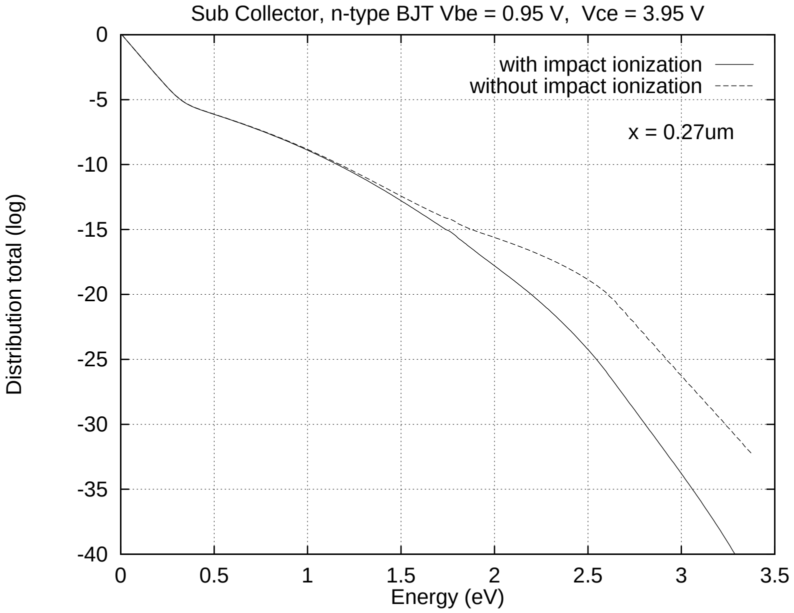


(b) Computed from a field-dependent model in drift-diffusion

Figure 6.10: Profile of impact ionization coefficients, α_{impact}, in the n-type and p-type BJT's computed by two methods: (a) solution of the Boltzmann transport equation by spherical harmonics; and (b) field-dependent ionization coefficient, using the Chynoweth formula. Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.



(a) In the collector



(b) In the sub-collector

Figure 6.11: Effect of including impact ionization on the distribution function. Total distribution (all bands) is plotted at two locations in the n-type BJT: (a) in the collector region; and (b) sub-collector region. The distribution is a weighted sum of distribution in bands (12) and (34); weighted by density of states. Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

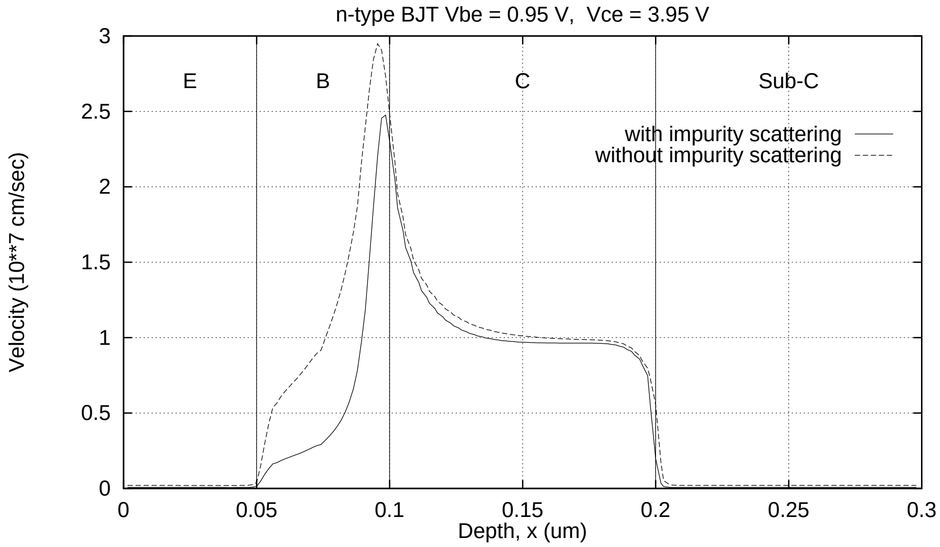


Figure 6.12: Effect of including ionized impurity scattering on the electron velocity. While there is a marked difference in velocity overshoot and velocity in neutral base, saturation velocity is unchanged. The n-type BJT is biased at $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

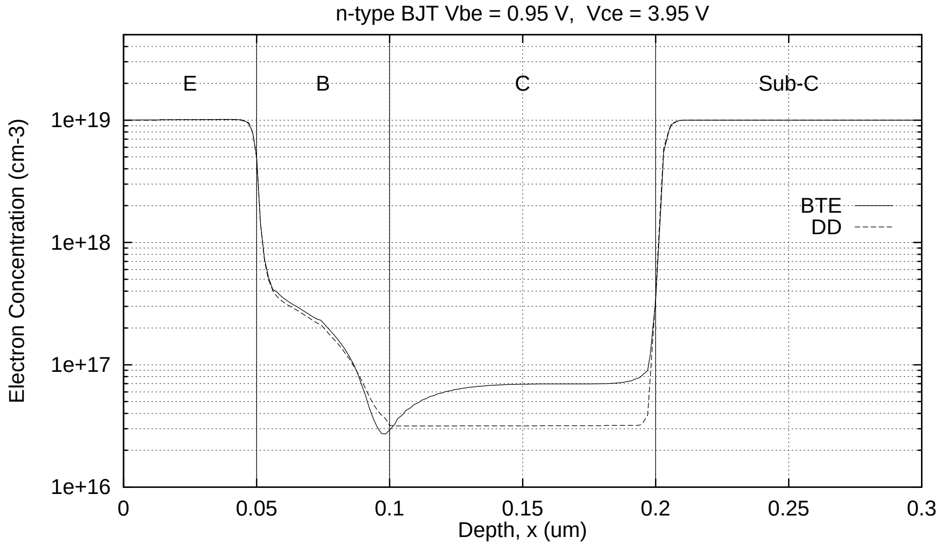


Figure 6.13: Comparison of electron concentration calculated by two techniques: spherical-harmonic Boltzmann transport equation (marked "BTE"); and drift-diffusion equations (marked "DD"). Bias in this n-type BJT is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

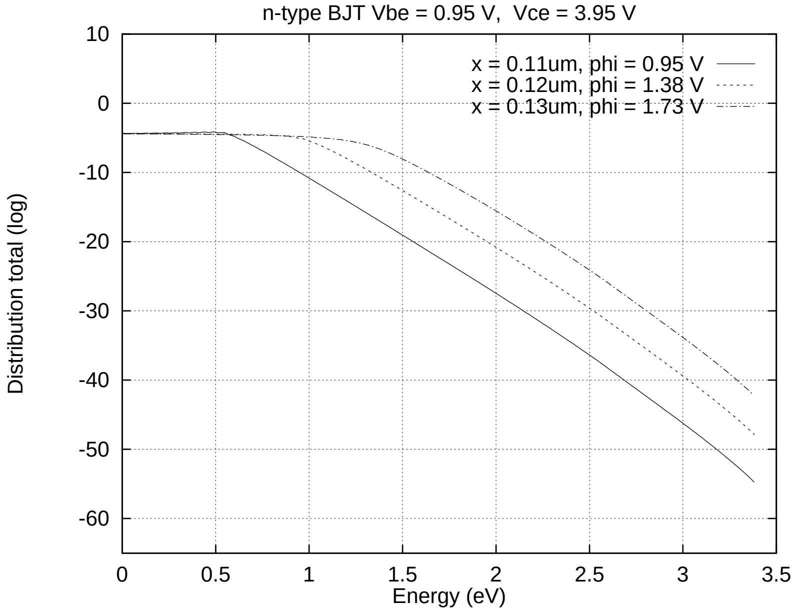


Figure 6.14: Ballistic transport in the n-type BJT. Distribution function is plotted at three locations, $x = 0.11$, 0.12, and 0.13μm; where the potential is $\phi(x) = 0.95$, 1.38, and 1.73V respectively. The distribution is a weighted sum of distribution in bands (12) and (34); weighted by density of states. Distributions are shifted from each other in energy by an amount $\Delta\varepsilon = q\,\Delta\phi$, indicating the presence of ballistic transport. Bias is $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

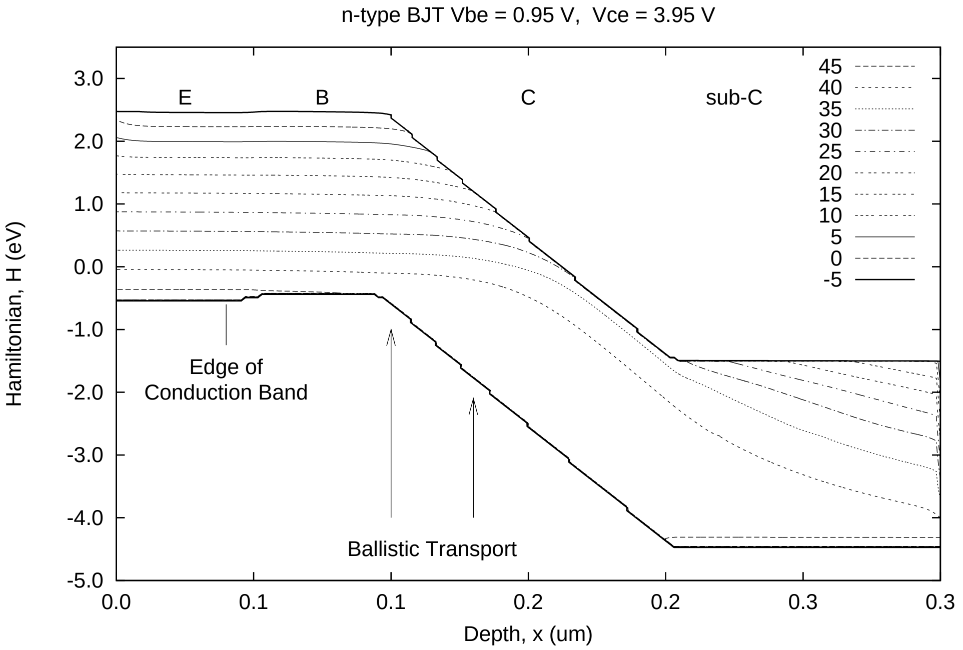
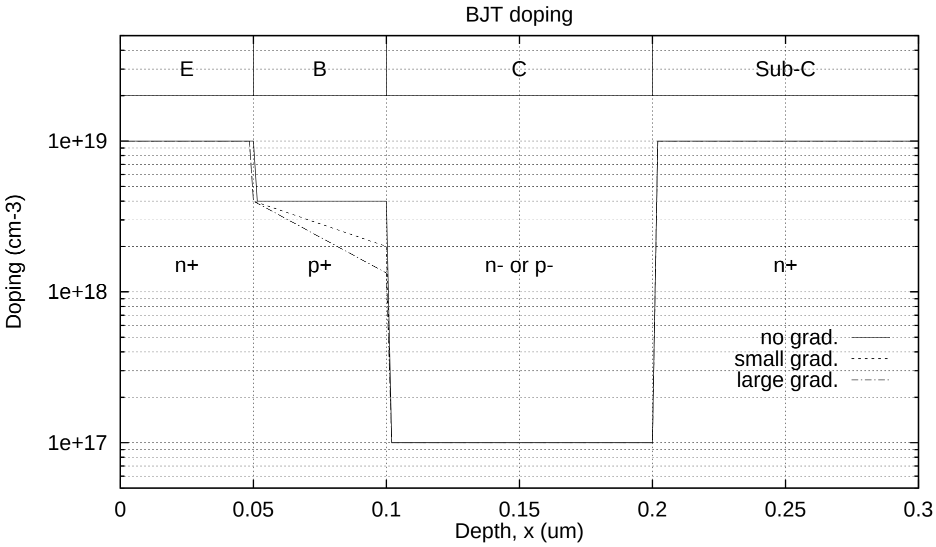
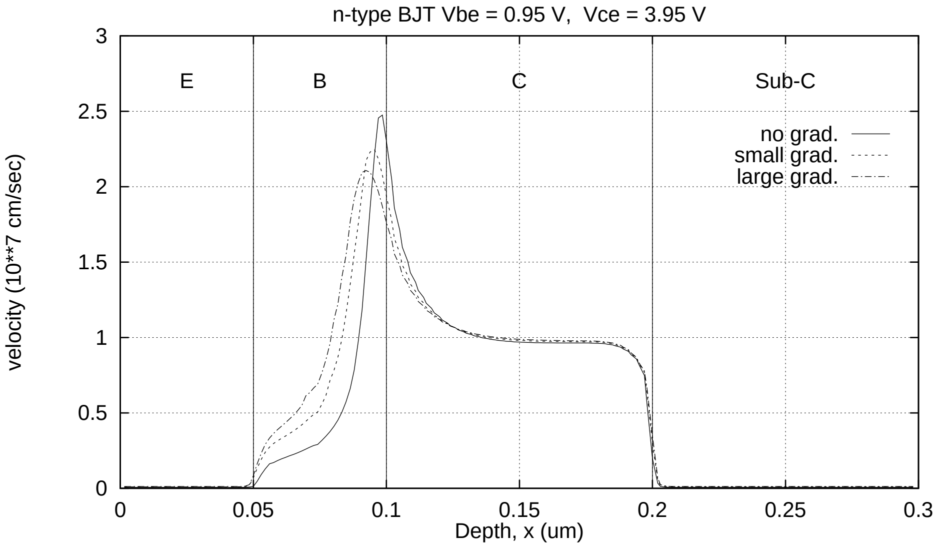


Figure 6.15: Ballistic transport in the base-collector space charge region of BJT. Contour plot of the electron density $\left(f_0^0(x,\varepsilon)g(\varepsilon)\right)$ in Hamiltonian, H, and space, x, for band (12) in n-type BJT. Units of the quantity being plotted has been purposely left unspecified; the only thing that *is* is important is that it is proportional to actual distribution of electrons in energy and space. Lower bound of contours is the edge of conduction band, $\varepsilon = 0$; and the upper bound is top of band (12), $\varepsilon = \varepsilon_{\text{max}}^{(12)}$. Bias is $V_{BE} = 0.95\text{V}$, and $V_{CE} = 3.95\text{V}$.



(a) Doping profiles in the base



(b) Electron velocity for various base doping profiles

Figure 6.16: Effect of graded base doping profile on electron velocity. Graded profile results in a built-in electric field which accelerates electrons. The n-type BJT was biased at $V_{BE} = 0.95$V, and $V_{CE} = 3.95$V.

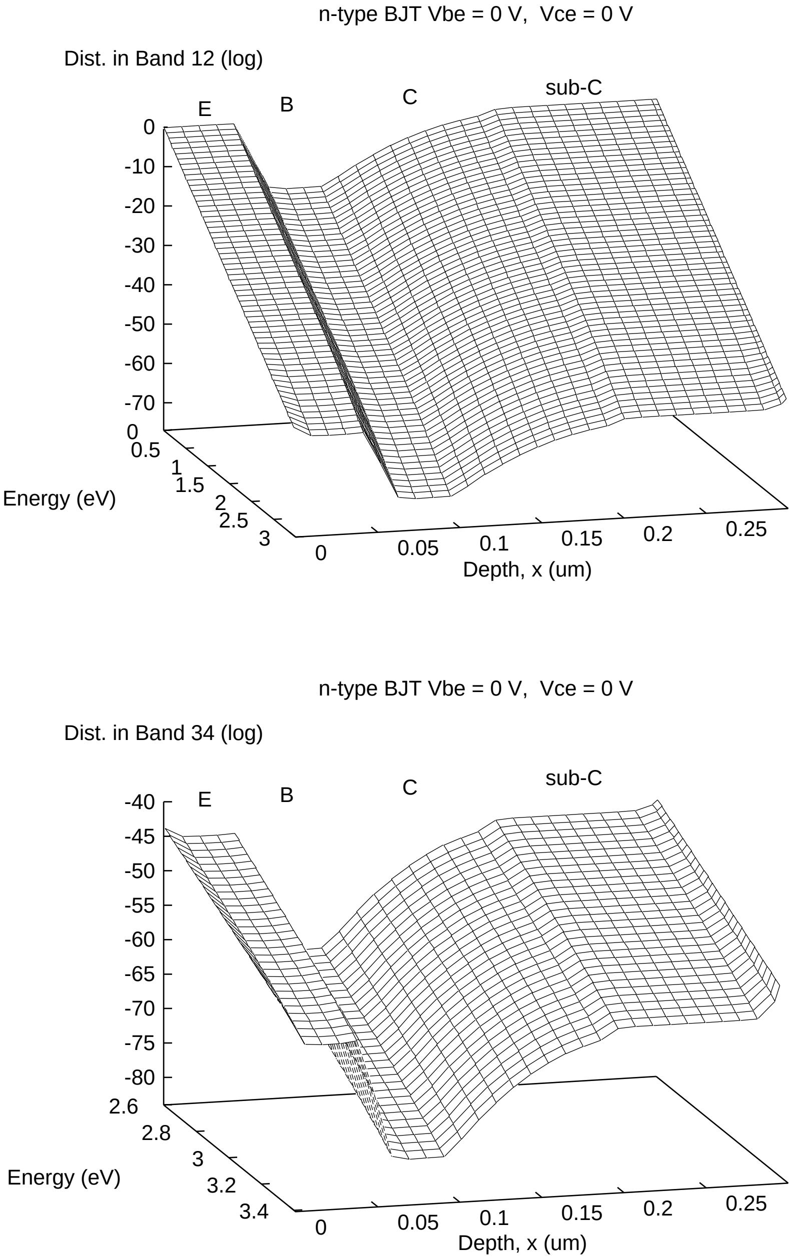
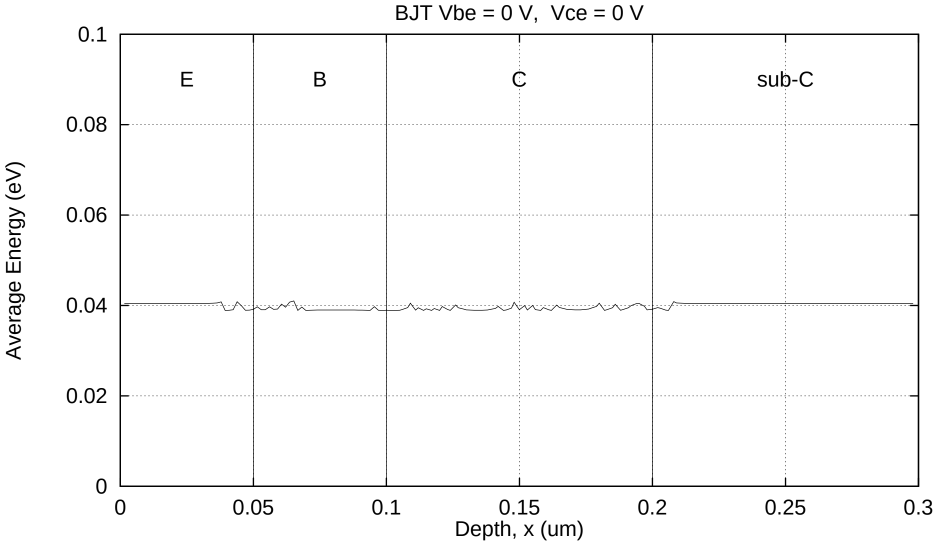
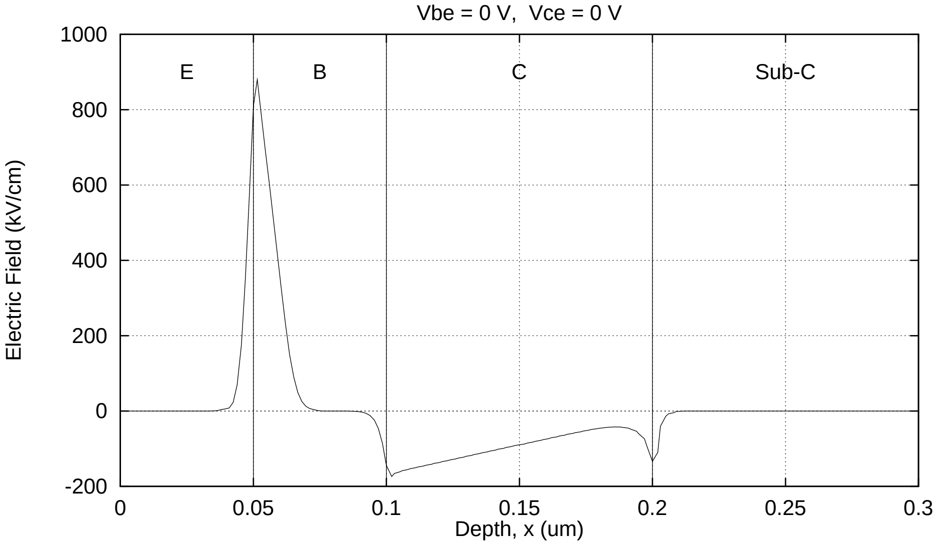


Figure 6.17: Distribution function in the n-type BJT at thermal equilibrium. All biases are zero at thermal equilibrium: $V_{BE} = 0\text{V}$, and $V_{CE} = 0\text{V}$. At thermal equilibrium the distribution in both bands, bands (12) and (34), is Maxwellian $(\propto \exp(-\varepsilon/k_B T_L))$.



(a) Average energy at thermal equilibrium



(b) Electric field at thermal equilibrium

Figure 6.18: Average electron energy and electric field in the n-type BJT at thermal equilibrium. The bias at thermal equilibrium is $V_{BE} = 0$V, and $V_{CE} = 0$V. Strong electric fields are present but the the distribution is still at equilibrium. The average energy is virtually constant at the room-temperature equilibrium value of $3/2\, k_B T_L = 0.0388$eV.

CHAPTER 7

SIMULATION OF MOSFET

7.1 Chapter Introduction

After developing the discretization for the spherical-harmonic Boltzmann equation in Chapter 4, and the Gummel's decoupled method for self-consistency in Chapter 5, we simulated n^+nn^+ device and bipolar junction transistor (BJT). Now we turn our attention in this chapter to the most important semiconductor device: Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET).

The MOSFET is, by any standard, the most ubiquitous and technologically-significant semiconductor device in human history. It is used for practically all digital circuits, from microprocessors to memories. In fact, MOSFET has become the measure for semiconductor technology itself: gate length of MOSFET in a technology, and the number of MOSFET's on a chip are a metric of semiconductor processes.

Traditional drift-diffusion equations for simulating semiconductors tend to loose validity for sub-micron (channel-length $< 1\mu$m) and deep sub-micron (channel-length $\ll 1\mu$m) MOSFET's [9,10]. For modern MOSFET's we need to model electron transport by the more-accurate Boltzmann transport equation, which can be solved either by the popular Monte Carlo method [52, 53, 62–65, 67,70,92], or the new spherical-harmonic approach [39,41–44,46,50,51]. Monte Carlo method is accurate and powerful, but has numerous disadvantages: it is computationally prohibitive; and has high statistical noise in the high-energy tail of the distribution, especially in high-doping low-field regions. Spherical-harmonic method, on the other hand, promises to be a good alternative; it is computationally efficient, and does not suffers from any statistical noise.

The first simulation of MOSFET's by spherical-harmonic approach in literature was by Gnudi, Ventura, and Baccarani in 1993 [39]. This was followed by [41–44, 46, 50, 51]. While simulations in [42–44, 51] were self-consistent, those in [39, 41, 46, 50] were not. All MOSFET simulations employed only the lower two bands of the multi-band band-structure from Brunetti *et al.* [57].

In this chapter we present the results of simulation for sub-micron MOSFET's. In contrast to previous simulations, all four bands from the multi-band band-structure [57] were used. Since a MOSFET is a two-dimensional device, the discretization developed in Chapter 4 has been extended to two space dimensions. Multi-band Boltzmann equations were cast in the pre-Maxwellian variables to enhance diagonal dominance and improve numerical properties. To keep the problem simple the Poisson and hole-continuity equation were dropped; the solution, therefore, is not self-consistent. While self-consistency is desirable, this approximation should not introduce significant error. It may be noted that now the Boltzmann equation has three dimensions—two in real space, one in energy—the complexity of this problem is comparable to a three-dimensional drift-diffusion simulation.

This chapter is organized as follows: Section 7.2 introduces the MOSFET device. The extensions of the previous numerical techniques are discussed in Section 7.3 on the facing page. In particular, this section introduces the boundary conditions used for MOSFET's, and the extension of the solver to two space dimensions. It also discusses the decoupled solution technique and grid generation. Section 7.4 on page 137 gives details of the MOSFET simulator. Results of the simulations are presented in Section 7.5 on page 138, where, amongst other issues, we analyze ballistic transport, velocity overshoot, thermal equilibrium, and effect of channel length in MOSFET's.

7.2 Introduction to MOSFET's

MOSFET is the most important semiconductor device invented. It is used for almost all digital, and many analog applications. MOSFET has gained a pre-eminent status in mainstream semiconductor industry and other semiconductor devices have been relegated to only highly-specialized applications.

The structure and principle of operation of a MOSFET is simple. Fig. 7.1 on page 146 shows a cross section through a MOSFET. First part of the name—metal-oxide-semiconductor, or MOS—underscores the structure of the device, we can see the gate terminal connected to the metal layer on top of the oxide, which resides on the semiconductor substrate. Second part of the name—field-effect transistor, or FET—emphasizes the principle of operation. Electric field generated by the voltage on the gate terminal results in transistor action. The word transistor itself comes from the word "trans-resistor," which refers to resistance modulation between two terminal due to electrical signal at some other terminal.

In addition we can see the *source*, *drain*, and *substrate* terminal. These terminals are ohmic contact to the device. The two elliptical, n^+-doped regions are the *source* and *drain* region of the MOSFET, with a junction depth of y_{junction}. The oxide has thickness t_{ox} and it resides on the lightly doped substrate. It overlaps the source and drain by an amount x_{overlap}. Amongst many measures of channel-lengths, two are shown in the figure: the metallurgical channel-length, $L_{\text{metallurgical}}$; and the gate length, L_{gate}.

7.3 Numerics for MOSFET Simulation

Simulation of MOSFET's is different from the previous simulations of n^+nn^+ device and bipolar junction transistors (BJT). Although most numerical techniques developed for those devices can be directly used for simulating MOSFET's, some require modifications. This section concentrates on giving a brief description of such modifications.

7.3.1 Boundary Conditions

Fig. 7.1 on page 146 shows the cross-section of a MOSFET. Boundary conditions from Section 2.7 on page 35 are applied at the various interfaces of the MOSFET.

Ohmic contacts to the semiconductor exist at the source (B-C), drain (D-E), and the substrate (A-F). At these contact we apply the usual ohmic Dirichlet boundary condition from Section 2.7.1.

Artificial boundaries (A-B & E-F) are introduced to isolate one MOSFET from rest of the chip for the purpose of simulation. This aim is best served by assuming zero current flux and applying the insulator boundary-conditions of Section 2.7.2.

Silicon-dioxide is one of the best insulators known to man. Therefore the semiconductor-oxide interface (C-D) is assumed to be insulating and we apply the insulator boundary-conditions of Section 2.7.2. It should be pointed out that some high-energy electrons can surmount the silicon-dioxide potential barrier and contribute to current across this interface. Such electrons are very few, and we can therefore choose to ignore their effect in our simulations.

7.3.2 Two Space Dimensions

The n^+nn^+ device and bipolar junction transistor were one-dimensional structures. Chapter 4 discussed the one-dimensional control-volume discretization which was applied to their simulation. But a MOSFET is essentially a two-dimensional device; in the channel region, for example, the electric field is orthogonal to electron flow direction. The spherical-harmonic Boltzmann equation in two space dimensions, after transforming to the pre-Maxwellian variables, is

$$\boxed{\frac{\partial}{\partial x}\left(\kappa\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial y}\left(\kappa\frac{\partial C}{\partial y}\right)+3g\left[\frac{\partial C}{\partial t}\right]_{\text{collision}}=0} \tag{7.1}$$

The control-volume discretization of this two-dimensional equation is a simple extension of the one-dimensional case and need not be discussed in detail.

7.3.3 Decoupled Solution Technique

The multi-band Boltzmann equations are solved as post-processor to the drift-diffusion solution: Electric potential from the solution of Poisson equation in drift-diffusion equations remains unchanged during the solution of the the Boltzmann equations. If the electron densities from the drift-diffusion model and the Boltzmann equation do not differ appreciably, we have incurred negligible error by having the solution as a post-processing.

The new flow-chart for the solution of the multi-band Boltzmann equation is shown in Fig. 7.2 on page 147. Solution of the drift-diffusion equations provides the electric potential $\phi(x,y)$ for the Boltzmann equations. Electron density from drift-diffusion, $n(x,y)$, is used to construct good initial guess for the variable $C(x,y,H)$; the distribution function is assumed to be Maxwellian $(C(x,y,H) = C_0(x,y))$ with the pre-Maxwellian factor chosen to yield the drift-diffusion electron density.

The two individual Boltzmann equations are solved by line successive over-relaxation (SOR) in the space and Hamiltonian direction. Boltzmann equation in one band is solved while keeping the other distribution from the other band constant. This procedure is repeated till distributions in both bands converges. The convergence criterion to stop iterations is the relative change of the $C(x_i,y_j,H_k)$ in every band:

$$\max_{\forall\, i,j,k}\left|\frac{C^{(m)}_{i,j,k}-C^{(m-1)}_{i,j,k}}{C^{(m)}_{i,j,k}}\right| < 10^{-7} \text{ to } 10^{-5} \tag{7.2}$$

where m is the iteration count. In the initial stages of solution of one band the distribution in the other band has not converged, therefore, we need not take the solution to full convergence; we limit the maximum inner loops to around thirty.

7.3.4 Grid Generation

The spherical-harmonic Boltzmann equation for MOSFET's, Eq. (7.1), is three-dimensional: two in real-space dimensions; and one in Hamiltonian (energy).

Discretization of this equation, therefore, can be result in large number of grid points. Large number of grid points result in more computation time, as well as require more computer memory (RAM). Therefore, we need to be to very frugal with grid points. We achieve this by making our grid nonuniform in space; in Hamiltonian (energy), unfortunately, we are forced to keep a uniform grid (refer to control-volume principle 4 on page 62). We allocate more grid points in the channel region and near the channel-drain and source-channel interface since the solution there varies rapidly; this can be seen in the grid shown in Fig. 7.3 on page 148 This is automatically done by the simulator in the following fashion.

First let us discuss the grid in the vertical, or y, direction. The MOSFET in vertical direction is divided into three regions: channel region, source-diffusion depth, and the remaining substrate. In each region the grid expands with a constant factor as the grid is specified from the top (semiconductor-oxide interface) towards the bottom (substrate contact):

$$\Delta y_{j-1} = r_e \, \Delta y_j \tag{7.3}$$

The constants r_e and the first Δy for each region are user-specified inputs to the program.

In the lateral, or x, direction a similar procedure first divides the MOSFET in four regions: source region; source-channel interface to the middle of the device; middle of the device to the channel-drain interface; and the drain region. In each region the user-specified r_e and Δx are used to expand or contract the grid gradually:

$$\Delta x_{i+1} = r_e \, \Delta x_i \qquad \text{or} \qquad \Delta x_{i-1} = r_e \, \Delta x_i \tag{7.4}$$

This has proved to be an efficient method to generate a smooth grid. Of course, some experimentation is required to give intelligent values for r_e, Δx, and Δy for each region. Three points need emphasis. Firstly, for the grid in vertical direction we need an extra-fine grid spacing close to the semiconductor-oxide interface: in fact, for the for the simulation we used $\Delta y = 4\text{Å}$ which expanded at the rate of $r_e = 1.15$. Secondly, in the lateral direction the channel-drain region, and the source-channel region to a lesser extent, need a higher grid density. This is because the electric field and doping varies more rapidly in this region. Thirdly, for simplicity the same grid is used for the drift-diffusion as well as the Boltzmann equation, this eliminates the need to interpolate.

7.4 The Simulator

The simulation program for the MOSFET, comprising of about 2,500 lines of uncommented FORTRAN 77 code, was run on a 200-MHz Pentium Pro personal computer (PC) with 64 MB of RAM (Random Access Memory) running a Linux RedHat 4.0 operating system using the gnu Fortran compiler. The drift-diffusion portion is external to the simulator, and was borrowed from a previous project [98]. Simulation time for the complete multi-band Boltzmann equation for one bias point was typically 30–45 minutes.

7.5 Results

This section presents some results for the simulation of sub-micron MOSFET's. A sub-micron MOSFET of effective channel-length of 0.17μm is analyzed in detail at various biases and thermal equilibrium. Following that, MOSFET's of varying channel-lengths are discussed.

7.5.1 Device Structure & Doping

A representative test-MOSFET was constructed by culling typical physical values from published sub-micron structures. The structure is realistic yet simple; the aim of this MOSFET is to serve as a test vehicle for the simulations, therefore, features like lightly-doped drain or channel implants have been ignored. Table 7.1 on page 145 lists the numerical values of various parameters. Effective channel-length of the MOSFET—which can be defined in a many ways; we define it as the metallurgical channel-length—is 0.17μm, while the gate length is 0.24μm.

Doping profile of the MOSFET is shown in Fig. 7.4 on page 149. In the source and drain region doping is constant in an elliptical region and then decays as a half-Gaussian, as $\exp(-(r-r_0)^2/2\sigma^2)$. The background substrate doping is constant. The gate is assumed to be heavily-doped n-type poly-silicon whose Fermi level is at the conduction-band level (refer to p. 206 of [81]).

7.5.2 Grid & Simulation Time

The grid for the MOSFET was generated by using the grid-generation approach described in Section 7.3.4 on page 136. This grid, as shown in Fig. 7.3 on page 148, has 38 points in the x-direction and 36 points in the vertical direction. Grid-spacing in energy is one-fifth of optical phonon energy jump: $\Delta H = \hbar\omega_{\mathrm{opt}}/5$. This resulted in a grid in energy that has 291 points in band (12), and 76 points in band (34). The total number of grid points are $38 \times 36 \times (291+76) \approx 500{,}000$. Solution of the multi-band Boltzmann equations on this grid usually takes around 30–45 minutes for one bias point.

7.5.3 I-V Curves

From the usual MOSFET theory, the *I*-*V* characteristics in the strong-inversion regime is [83]:

$$I_{DS} = \frac{\mu_n C_{\text{ox}} W}{L}\left(\left(V_{GS}-V_T\right)V_{DS}-\frac{V_{DS}^2}{2}\right) \qquad \text{For} \quad \begin{array}{l} V_{GS} > V_T \;\; \& \\ V_{DS} < V_{GS}-V_T \end{array} \tag{7.5}$$

$$I_{DS} = \frac{\mu_n C_{\text{ox}} W}{2L}\left(V_{GS}-V_T\right)^2 \qquad \text{For} \quad \begin{array}{l} V_{GS} > V_T \;\; \& \\ V_{DS} > V_{GS}-V_T \end{array} \tag{7.6}$$

and in the sub-threshold regime it is

$$I_{DS} = \frac{W}{L} I_{D0} \exp\left(\frac{qV_{GS}}{nk_BT_L}\right) \qquad \text{For} \quad \begin{array}{l} V_{GS} < V_T \;\; \& \\ V_{DS} > 3\,k_BT_L/q \end{array} \tag{7.7}$$

where I_{DS} is the drain-to-source current; V_{DS} is the drain-to-source voltage; V_{GS} is the gate-to-source voltage; and V_T is the threshold voltage; W is the width and L is the effective channel length of the MOSFET. Other constants need not be explained. We can use these simple idealized equations to compare the simulated *I*-*V* curves.

The plot of the simulated drain-to-source current I_{DS} against the drain-to-source voltage V_{GS}, also known as the transfer characteristics, is shown in Fig. 7.5 on page 150. In Fig. 7.5(a) the drain current is plotted at a constant drain-to-source voltage of $V_{DS} = 50$mV. This is a typical value chosen to infer the threshold voltage (p. 438, [84]). For gate voltage larger than 1V the curve is seen to be linear, as expected from Eq. (7.5) for small drain voltage. From the intersection of this straight line with the gate-voltage axis we can infer a threshold voltage $V_T \approx 0.7$V. This method of threshold-voltage extraction is called *linear extrapolation method* (p. 439, [84]).

In the sub-threshold regime the simple MOSFET theory predicts the drain current to depend exponentially on the gate voltage, as seen in Eq. (7.7). This is seen in Fig. 7.5 where the *I*-*V* curve is a straight line for $V_{GS} < V_T$.

Fig. 7.6 on page 151 plots the simulated output characteristics of the MOSFET. It displays features that are typical to a MOSFET *I*-*V* curve: a linear region for low drain voltages; a saturation region for high drain voltages; and a transition to saturation at the drain voltage of approximately $V_{DS,\text{sat}} = V_{GS} - V_T$.

Substrate current is produced due to holes generated by impact ionization (Chapter 8, [84]; and [85]). In fact, substrate currents are used as a measure of hot-electron effects in a MOSFET (Chapter 8, [84]; and p. 29, [85]).

The simulator estimates substrate current assuming that all generated holes drift to the substrate contact. Fig. 7.7 on page 151 shows the substrate current versus gate voltage, for various drain voltages; this is a customary way of plotting substrate current (p. 370, [84]). For a given drain voltage, substrate current initially increases with increasing gate voltage because of increasing drain current. Further increase in gate voltage results in lower substrate current; this is due to a reduction of electric field. The maximum occurs at about $V_{GS} \approx (0.3 \text{ to } 0.5)\, V_{DS}$ (p. 370, [84]).

Note On Current calculations

Since we used the current-conserving control-volume discretization for the Boltzmann equation, the current at the source contact is equal to current at the drain ohmic. In addition, this current is equal to current crossing a vertical cross-section through the MOSFET oxide. Fig. 7.8 on page 152 shows the current crossing through a vertical cross-section. Current is constant when the cross-section passes through the oxide. The current, therefore, is the same regardless of how we calculate it.

7.5.4 Distribution Function

To present the simulation results we apply the following voltage: Gate voltage, V_{GS}, is set to 1.5V; drain voltage, V_{DS}, is set to 2.5V. The drain voltage is sufficiently high to show hot-electron effect. Electric potential calculates from the drift-diffusion solution at this bias is shown in Fig. 7.9 on page 152.

Vector plot of the electric field is shown in Fig. 7.10 on page 153. Even though the electric field is vertical in the channel region, the transport occurs mainly in the lateral direction. The device has high electric fields that vary rapidly in space. This is due to both the applied bias, as well as the built-in electric fields originating from high doping concentrations.

The multi-band distribution function is presented here at three cross-sections in the MOSFET. The cross-sections are horizontal, parallel to the channel, at three levels inside the device: (a) just below the semiconductor-oxide interface; (b) at the level of the source/drain junction-depths; and (c) at a level deep into the substrates, away from the active region.

Fig. 7.11 on page 154 plots the distribution very near the semiconductor-oxide interface. Electron are move from the source to the drain. Distribution is Maxwellian ($\propto \exp(-\varepsilon/k_B T_L)$) in the heavily-doped field-free source region. As the electrons enter the channel region they experience mild lateral electric field, and gain some velocity. Nearer to the drain region is high, electric field is high; electrons gain energy and heat-up considerably and scatter to the upper bands. This can be seen by the non-Maxwellian shape of the distribution near the drain region. We also see the distribution of band (34) increase by many orders of magnitude. Finally electrons in the heavily-doped field-free drain the distribution tends to relax to the Maxwellian shape. In fact, based on the slope of the distribution in the drain, we can identify two types of electron: sea of cool equilibrium Maxwellian electrons in the drain, as seen by the slope at low energies; and a small population of hot high-temperature electrons injected from the channel, as seen by the slope at higher energies [110].

Fig. 7.12 on page 155 plots the distribution along the cross-section at the source/drain junction-depth. Although electrons at this level are not as heated as those near the interface, we can see a little non-Maxwellian shape near the drain region, which is actually the lower edge of the drain region. Distribution function deep into the substrate, as shown in Fig. 7.13 on page 156, is expected to be equilibrium Maxwellian because there is negligible electron current and almost no electric field.

7.5.5 Average Quantities

Fig. 7.14 on page 157 displays the current density in the MOSFET as calculated by two methods: drift-diffusion and Boltzmann transport equation. One principal difference to notice between the two is the *spreading* of the electrons near the drain in the Boltzmann equation method. This is not surprising; it has been explained by Meinerzhagen [111] as *thermal diffusion.* Electrons at the drain are hot, hence, have more tendency to diffuse; this results in a higher electron concentration than one might get from drift-diffusion [52, 111]. Semiconductor models which account for electron energy, like hydrodynamic, energy balance, or Monte Carlo, encounter this “thermal difusion” spreading: refer to Fig. 17 & 19 in [93]; or Fig. 8.5 in [55]; and [52, 111].

Fig. 7.15 on page 158 shows the current density in the MOSFET. Current is conducted mainly in a very thin channel region. Fig. 7.16 on page 159 displays the average electron velocity in the device. Surface plot of average electron energy is shown in Fig. 7.17 on page 160. While electrons in most of the device are at equilibrium, electrons near the drain region possess high energies. These hot electrons cause impact ionization, as shown in Fig. 7.18 on page 161. We see that impact ionization is negligible in the rest of the device, but extremely high near the drain. The holes generated by impact ionization cause undesirable substrate currents.

From the previous discussion it is clear that the active region of the MOSFET is the channel created directly under the gate oxide. One-dimensional plots of quantities in the channel would make the presentation more lucid. Fig. 7.19 on page 162, and Fig. 7.20 on page 163 plot some quantities in the channel. We can see Fig. 7.19 that most of the potential drops near the drain, therefore, the electric field is high near the drain. Average energy follows the electric field, although its peak is somewhat beyond the maximum of the electric field. This delay is known as the *dead-space effect* [94], or sometimes also as the *dark-space effect* [69]. In Fig. 7.20 on page 163 we see that electrons that enter this high-electric-field area overshoot their saturation velocity of 10^7 cm/s. The high energy electrons also impact ionize, as shown in Fig. 7.20; the peak of impact ionization is located beyond the maximum of either electron energy or electric field. This can be explained as follows. Even though there are substantial number of hot electrons, average energy is low because of the sea of cool electrons. Impact ionization depends only on the tail of the distribution function; average energy, however, is more sensitive to the distribution function at low energies.

Fig. 7.21 on page 164 displays the electron concentration in bands (1), (2), (3), and (4) under the oxide. In the source region the distribution is equilibrium Maxwellian; the upper bands, therefore, are sparsely populated. As electrons travel through the channel, electron concentration in the upper bands increases drastically due to both electric field and inter-band scattering. These hot electrons are a serious reliability concern for small-channel MOSFET's. In the highly-doped low-field region of the drain, electrons in higher bands begins to relax to the lower bands by inter-band scattering. If we included a very long drain in the simulation, the distribution would approach equilibrium Maxwellian.

7.5.6 Ballistic Transport

It was shown in Section 6.4.13 on page 114 that when electrons encounter large electric fields they may travel *ballistically*, or without scattering. It was also shown that the distribution function has signature in ballistic transport: The distribution of electrons traveling ballistically simply shifts in energy by an amount equal to the energy gained from the field. If an electron travels ballistically for a distance Δx during which the potential changes by $\Delta\phi$ then the distribution function at $x_0 + \Delta x$ is simply a shifted version of distribution at x_0:

$$f_0^0(x_0 + \Delta x, \varepsilon + \Delta\varepsilon) = f_0^0(x_0 + \Delta x, \varepsilon + q\,\Delta\phi) = f_0^0(x_0, \varepsilon) \qquad (7.8)$$

With this in mind we can view the distribution at three locations in the MOSFET channel just under the oxide in Fig. 7.22 on page 165. Similar to the discussion in Section 6.4.13 on ballistic transport in BJTs, we see that the distribution has two distinct slopes: a tail region in equilibrium with the lattice ($\propto \exp(-\varepsilon/k_B T_L)$); a low-energy region at a very high effective temperature (almost flat in energy). There is a distinct breakpoint between the two regions. The distribution at 0.15μm, however, has a negligible high-temperature region; it is almost entirely at equilibrium. Based on the position of the breakpoint in the distributions, we see that the distributions are shifted in energy by an amount $\Delta\varepsilon = q\,\Delta\phi$ from one another. This points to the existence of ballistic transport.

7.5.7 Thermal Equilibrium

The distribution function at thermal equilibrium, as proved in Appendix B on page 177, is Maxwellian. Thermal equilibrium is a good test for the numerics and the code, as well an interesting phenomena by itself. It was applied in Section 5.6 on page 89 for the $\mathrm{n^+nn^+}$ device, and in Section 6.4.15 on page 115 to the BJT. In this section we examine the MOSFET at thermal equilibrium.

Thermal equilibrium is defined as a state of no currents in the device. In a MOSFET that we need to set the drain voltage V_{DS} to zero; the gate voltage V_{GS} can be set to anything since it does not cause any current, but we will set to zero also for simplicity. Fig. 7.23 on page 166 displays the distribution function in the MOSFET at thermal equilibrium. We can see that the distribution function in both bands is Maxwellian at lattice temperature (it is $\propto\ \exp(-\varepsilon/k_BT_L)$).

The distribution is a Maxwellian despite the presence of large electric fields, as seen in Fig. 7.24(a). The Maxwellian nature of the distribution is further proven by the plot of the average energy in the device in Fig. 7.24(b); it is virtually constant at the thermal-equilibrium value. The average energy of a Maxwellian distribution is $3/2\,k_BT_L$ (p. 106 of [74]), which at room temperature is approximately 0.0388eV.

Electron velocity at thermal equilibrium, not plotted, was found to be virtually zero throughout the device. This, as pointed out in Section 5.6 on page 89, can be interpreted as *velocity undershoot* or *velocity damping*: electron velocity is much smaller than what is dictated by the local electric field.

7.5.8 Effect of Channel-Length

Fig. 7.25 on page 168 displays the effect of channel-length on MOSFET operation, where average quantities in the channel of four MOSFET's of effective channel-lengths 0.05, 0.15, 0.25, and 0.35μm are displayed. Applied bias is same for all MOSFET's. The maximum electric field at all lengths is quite similar; this is because most of the potential drops near the drain. The electric field profile in the 0.35-μm MOSFET is typical of a long-channel MOSFET: negligible electric field in the channel, most of it appearing near the drain. Velocity overshoots its saturation value of 10^7cm/s for all channel-lengths. The average velocity in the long-channel 0.35-μm MOSFET is small in the channel, but overshoots in the drain depletion region. For the short-channel 0.05-μm MOSFET, most of the channel is under the velocity overshoot region. This underscores the importance of non-equilibrium transport, which can model velocity overshoot, in device operation.

Transconductance of a MOSFET, which is defined as

$$g_m = \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{DS}=\text{Constant}} \tag{7.9}$$

is plotted in Fig. 7.26 on page 169 for the four channel-lengths. As expected, it increases for smaller channel-lengths.

7.6 Chapter Summary

In this chapter the spherical-harmonic Boltzmann-equation solver was extended to two space dimensions. This two-dimensional solver is comparable to a three-dimensional drift-diffusion solver in complexity. The simulator was applied to simulate MOSFET's in a post processing fashion; potential was assumed to be same as that obtained by drift-diffusion. A sub-micron MOSFET of channel-length of 0.17μm was analyzed in detail. Various non-equilibrium non-local effects were studied. It was shown that there is ballistic transport and velocity in the channel. I-V curves were plotted and compared to the usual MOSFET theory to demonstrate the capabilities of the simulator. Four MOSFET's, with channel-lengths down to deep sub-micron regime, were simulated. Velocity overshoot was found to be an important element in MOSFET's, especially at shorter channel-lengths. The case of thermal equilibrium in a MOSFET was analyzed, it revealed equilibrium conditions at high built-in electric fields.

Parameter	Value	Comment
L_{gate}	0.24μm	Gate length
$L_{\text{effective}}$	0.17μm	Metallurgical channel-length
$N_{D,\ \text{peak}}$	10^{20} cm^{-3}	Peak doping of Gaussian in source and drain
σ	0.28μm	Standard deviation of Gaussian doping
N_A	6.5×10^{17} cm^{-3}	p-type Substrate doping
t_{ox}	50Å	Oxide Thickness
V_T	0.7V	Threshold Voltage
x_{overlap}	0.035μm	Oxide overlapping n$^+$ diffusion
y_{junction}	0.025μm	Junction depth

Table 7.1: Numerical values of the various parameters for the simulated MOSFET. Physical meaning of these parameters is explained in Fig. 7.1 on the next page.

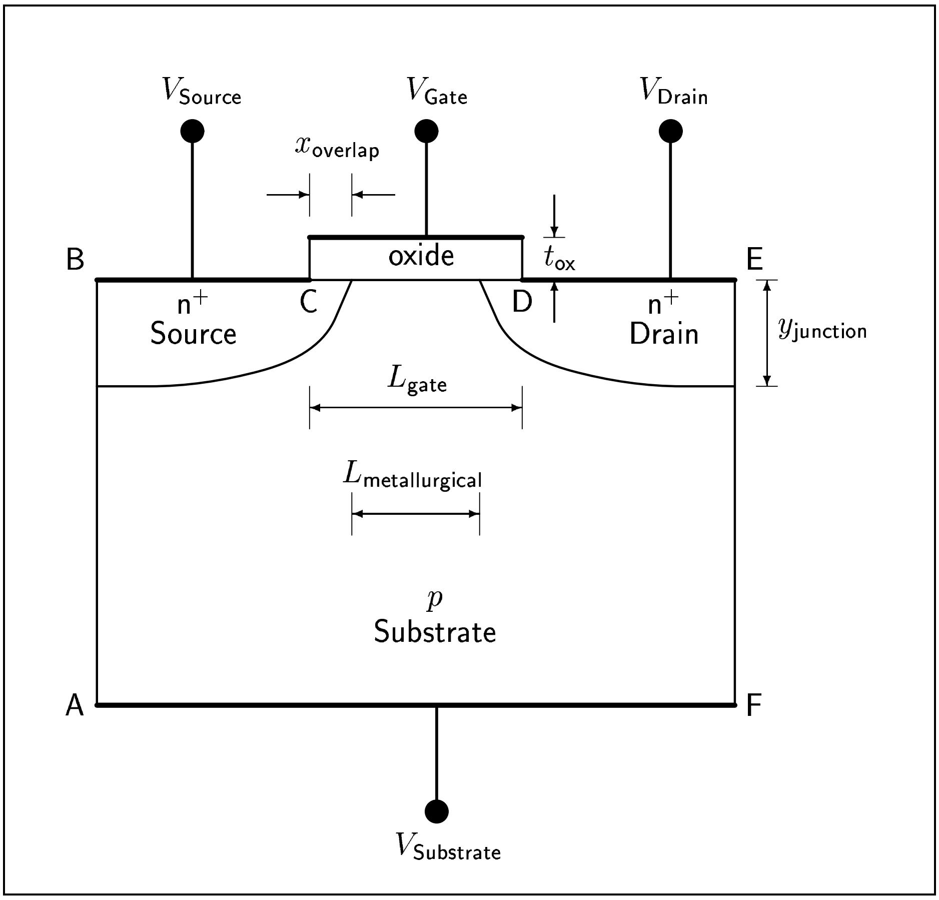

V_{Source}
V_{Gate}
V_{Drain}
x_{overlap}
oxide
t_{ox}
B
E
C
D
n$^+$ Source
n$^+$ Drain
y_{junction}
L_{gate}
$L_{\text{metallurgical}}$
p Substrate
A
F
$V_{\text{Substrate}}$

Figure 7.1: Schematic of a conventional n-channel MOSFET. n$^+$ source and drain, and p substrate is shown. B-C, D-E, and A-F are ohmic contacts. Channel length can be specified as either the metallurgical or the and gate length. Oxide thickness is t_{ox}; source/drain junction diffusion depth is y_{junction}; and overlap of source/drain with oxide is x_{overlap}.

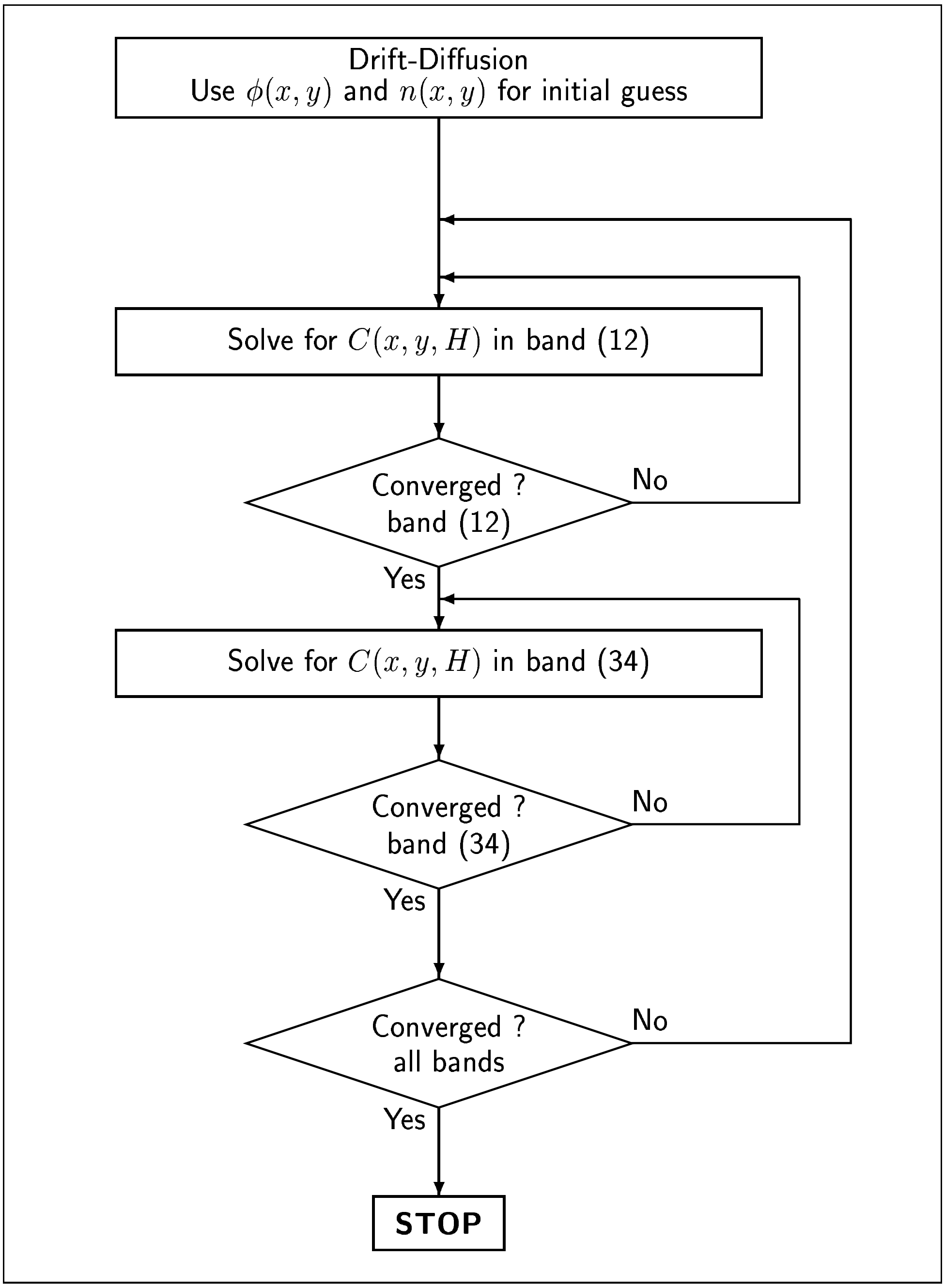

Drift-Diffusion
Use $\phi(x, y)$ and $n(x, y)$ for initial guess
Solve for $C(x, y, H)$ in band (12)
Converged ? band (12)
No
Yes
Solve for $C(x, y, H)$ in band (34)
Converged ? band (34)
No
Yes
Converged ? all bands
No
Yes
STOP

Figure 7.2: Flow chart for MOSFET simulation. Drift-diffusion simulation provides the electric potential, $\phi(x, y)$, and the initial electron concentration, $n(x, y)$. Boltzmann equation is solved iteratively in band (12) and (34).

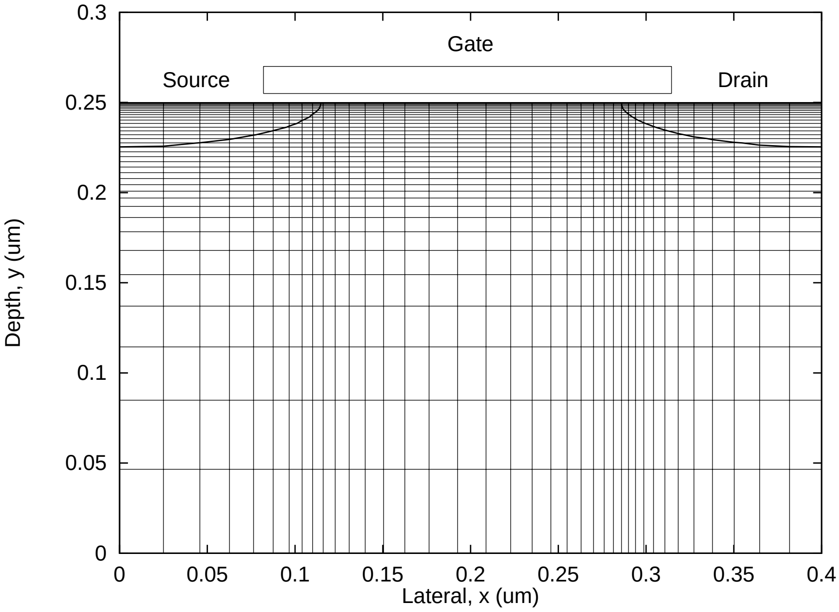


Figure 7.3: Grid in the simulated MOSFET. Lateral direction, x, is along the channel from the source to the drain. Depth, y, is in the direction perpendicular to the channel. Boundary of the *Source* and *Drain* is demarcated by heavy lines. At this boundary net doping is zero: $N_D = N_A$. Number of grid points in lateral direction N_x = 38, and in the vertical direction N_y = 36. Bands (12) and (34) have 291 and 76 grid points respectively; and $\Delta H = \hbar\omega_{\text{opt}}/5$. Total number of grid points are $N_x \times N_y \times (N_H^{(12)} + N_H^{(34)}) \approx$ 500,000.

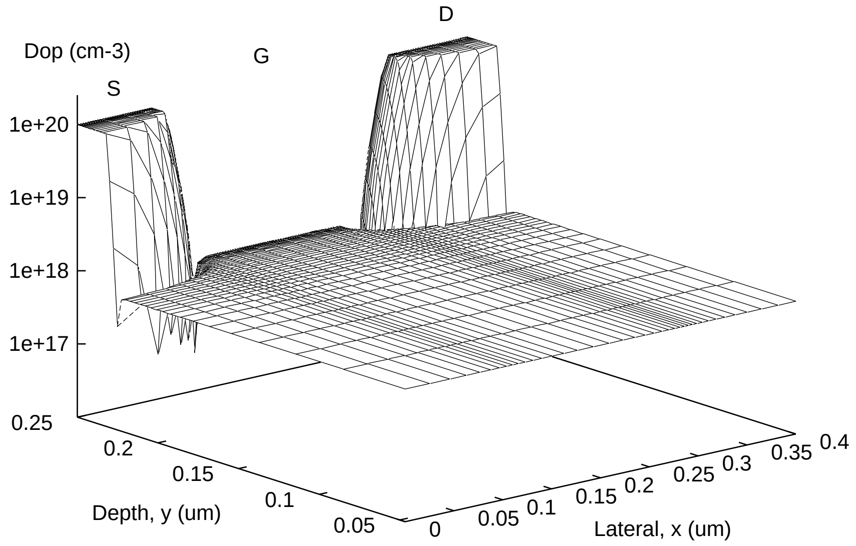
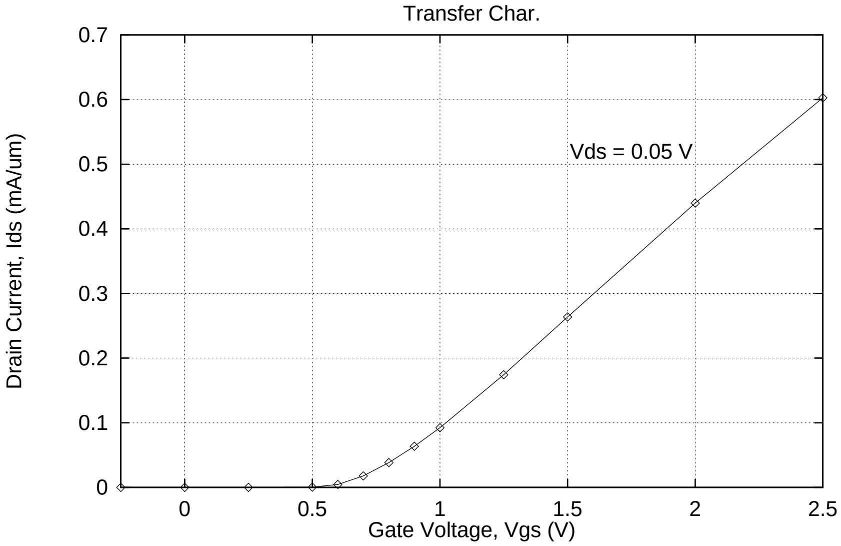
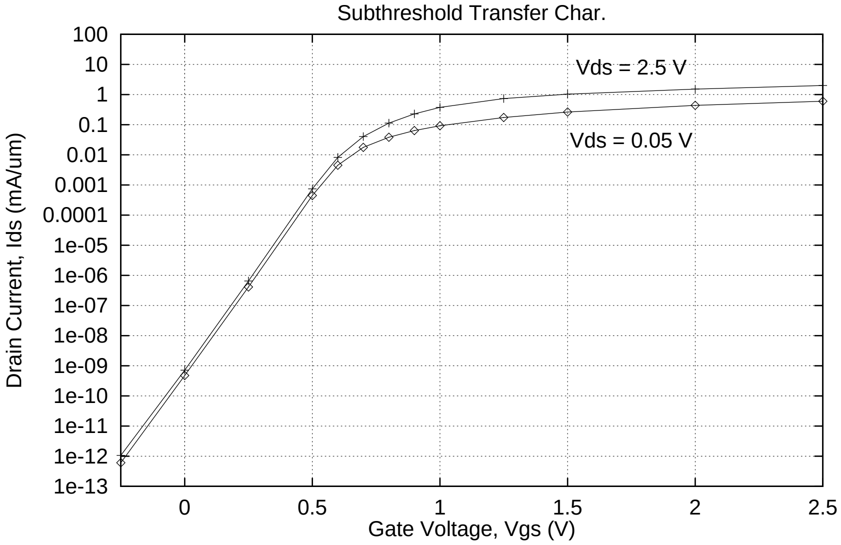


Figure 7.4: Doping profile of the simulated MOSFET. Plotted doping concentration is the net doping, difference between donor- and acceptor-type doping: $\mathcal{D} = N_D - N_A$. Source, gate, and drain are marked as "S," "G," and "D" respectively. Doping near the source and drain is n-type, while it is p-type in the rest of the device. Metallurgical channel-length is 0.17μm. The source and drain doping is constant in an elliptical region, then decays as a Gaussian, $\exp(-(r-r_0)^2/2\sigma^2)$.



(a) Transfer Characteristics



(b) Sub-threshold transfer Characteristics

Figure 7.5: The simulated transfer characteristics of the MOSFET. The transfer characteristic is the drain-current-versus-gate-voltage (I_{DS}-V_{GS}) characteristic of the MOSFET. Sub-figure (a) plots the transfer characteristics at a drain voltage V_{DS} = 50mV on a linear scale. From this curve we infer the threshold voltage, $V_T \approx$ 0.7V. In sub-figure (b) the sub-threshold transfer characteristics are plotted for V_{DS} = 50mV and 2.5V on a logarithmic scale. For $V_{GS} < V_T$, the *I-V* curve is seen to be linear on a log scale.

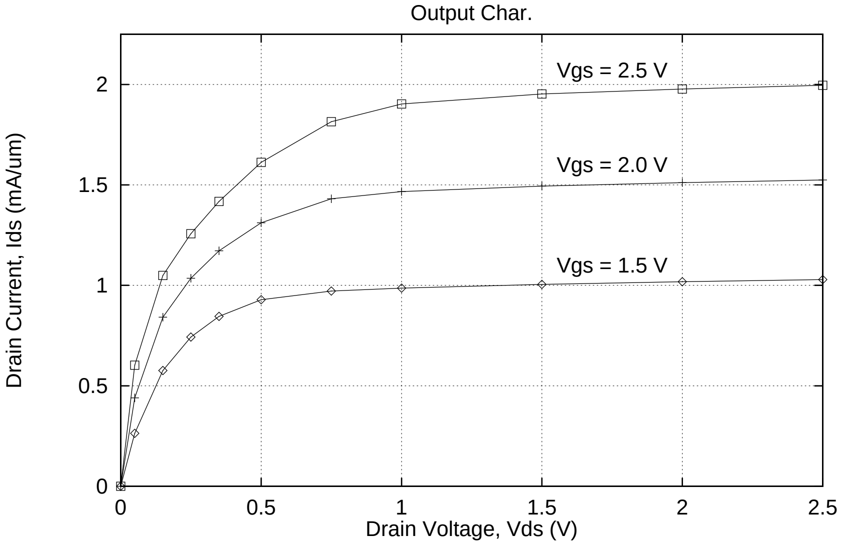


Figure 7.6: The simulated output characteristics of the MOSFET. Drain current I_{DS} is plotted against drain voltage V_{DS} at gate voltages V_{GS} = 1.5, 2.0, and 2.5V.

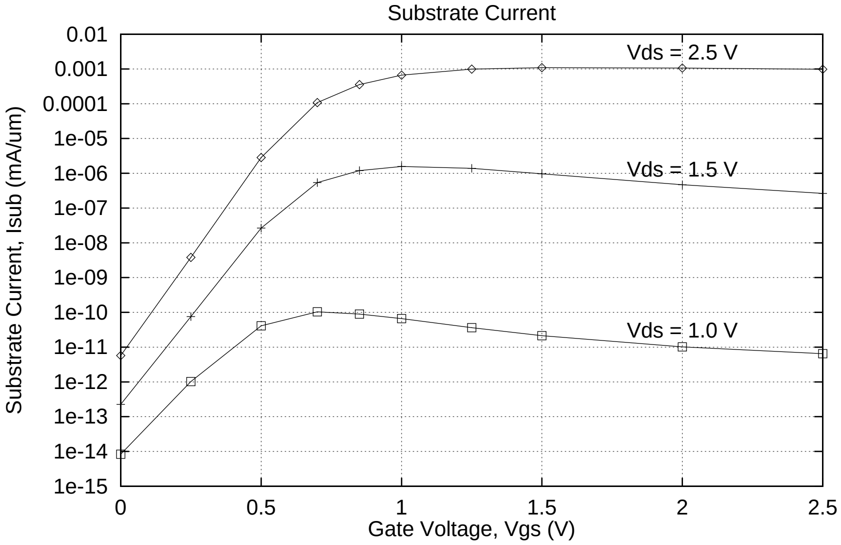


Figure 7.7: Simulated substrate current as function of gate voltage, V_{GS}, for various drain voltages, V_{DS}, in the MOSFET.

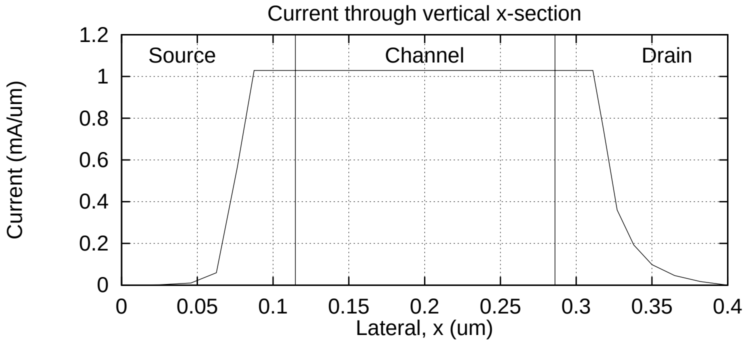


Figure 7.8: Current crossing through a vertical cross-section in the MOSFET. the current is constant when the vertical cross-section passes through the oxide. The figure demonstrates the current-conserving property of the control-volume formulation. The current in the center is equal to the terminal drain current, which in turn is equal to the terminal source current. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

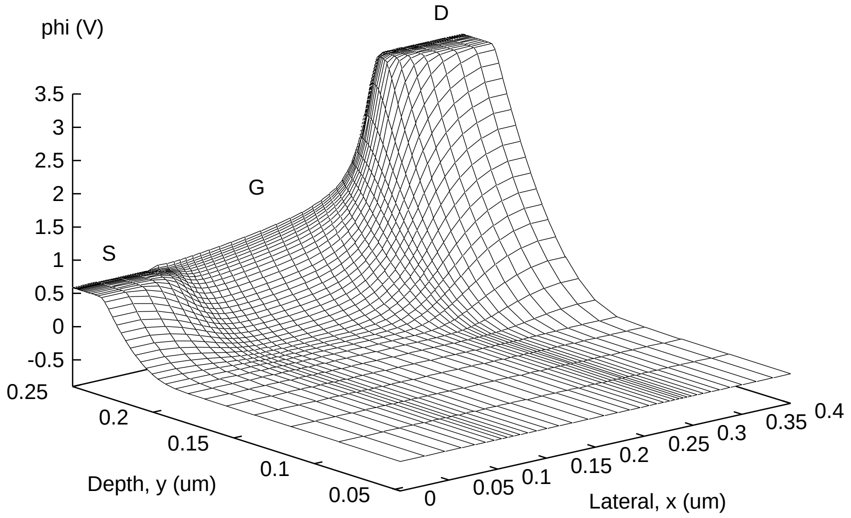


Figure 7.9: Electric potential, $\phi(x, y)$, within the MOSFET. Source, gate, and drain are marked as "S," "G," and "D" respectively. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

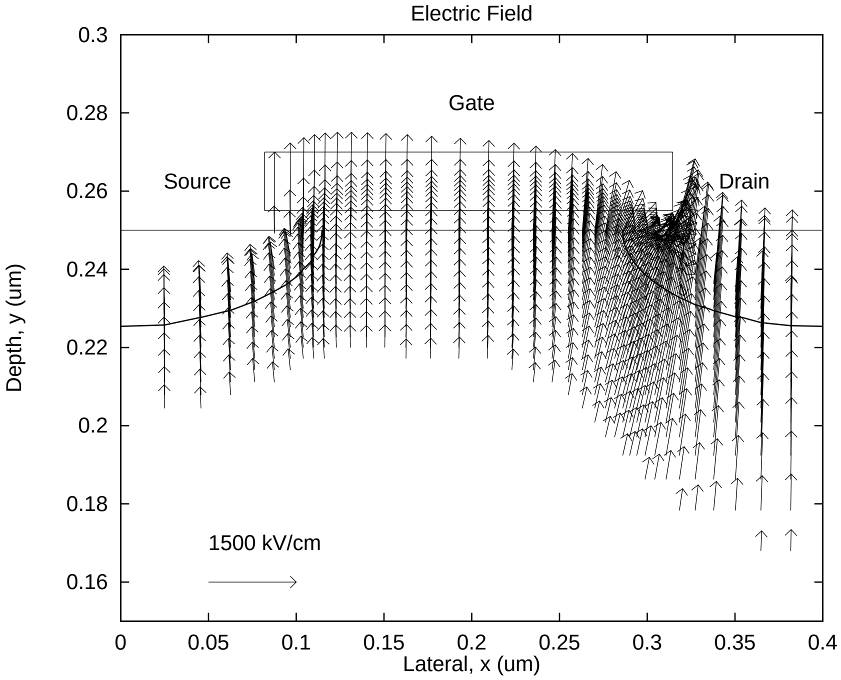


Figure 7.10: Vector plot of the negative of electric field within the MOSFET. Plotting the negative of the electric field indicates the direction in which the electron is experiencing force due to the electric field. Scale is shown in the lower left corner. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

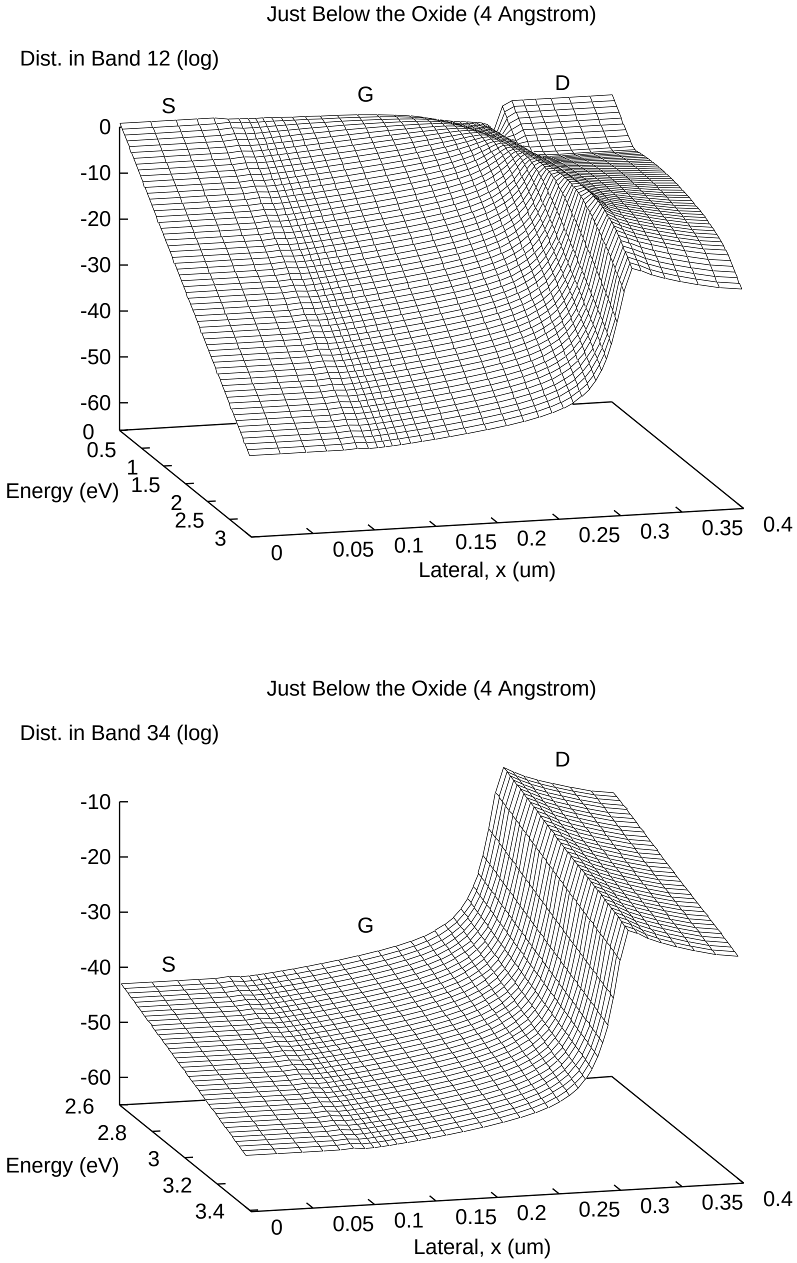
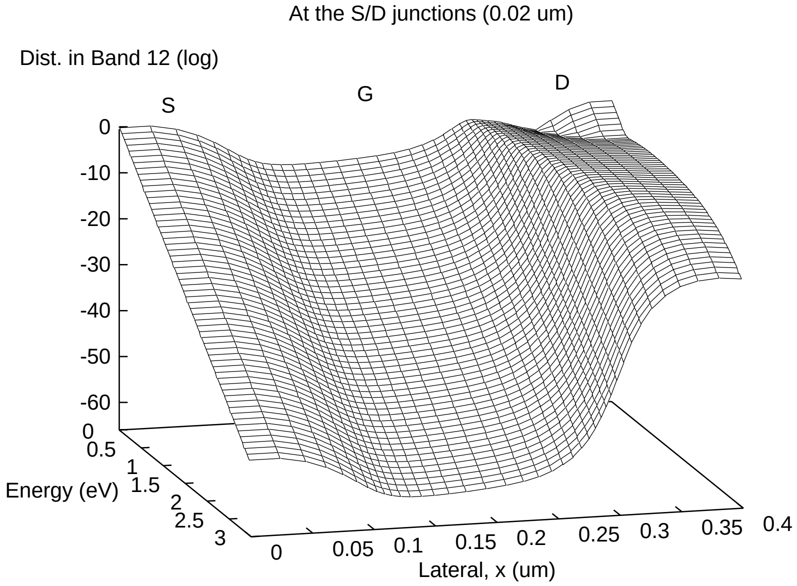


Figure 7.11: Distribution function along a horizontal section of the MOSFET under the oxide at the depth of 4Å. This is actually at the very first grid line under the oxide. Top figure is for the lower band, band (12); and the lower figure is for the higher bands, band (34). Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.



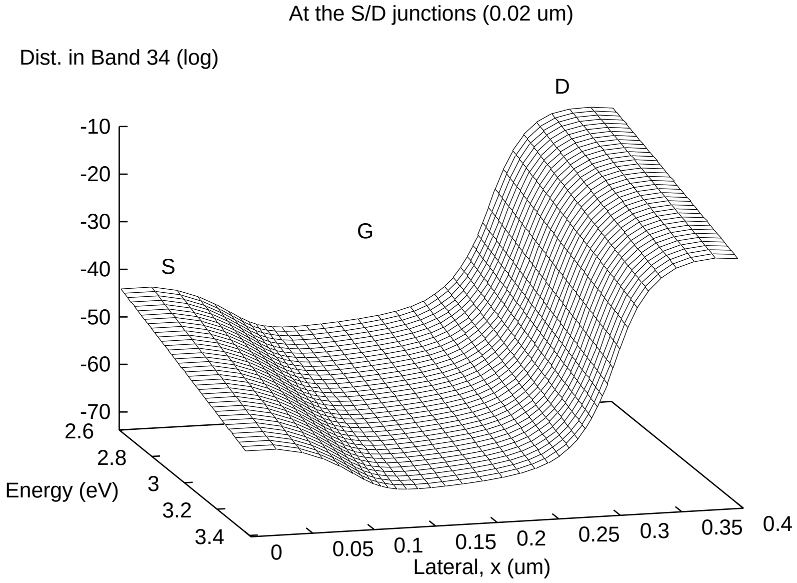
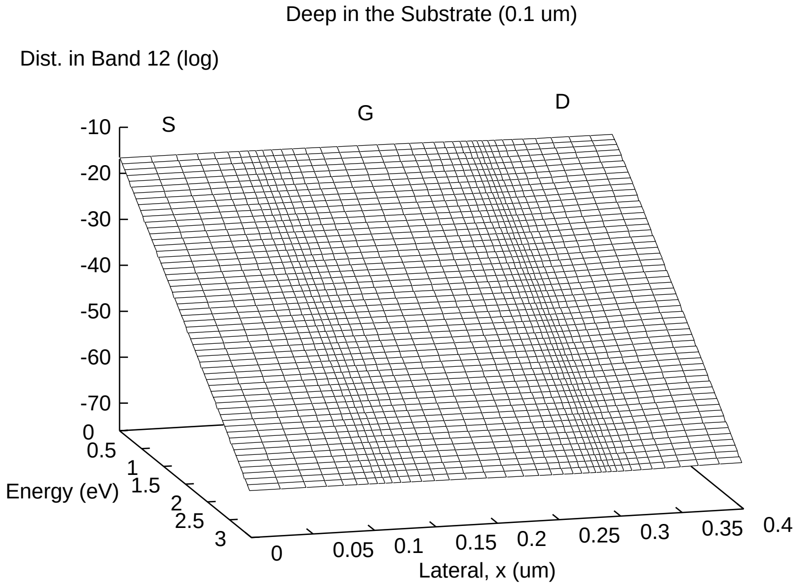


Figure 7.12: Distribution function along a horizontal section that is 0.02μm below the semiconductor-oxide interface, which is approximately the same as source/drain junction depths. Top figure is for the lower band, band (12); and the lower figure is for the higher bands, band (34). Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.



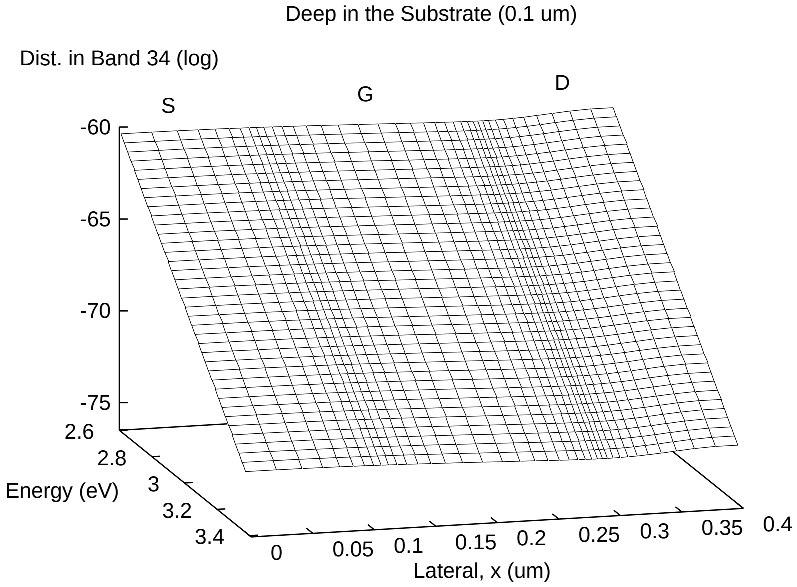
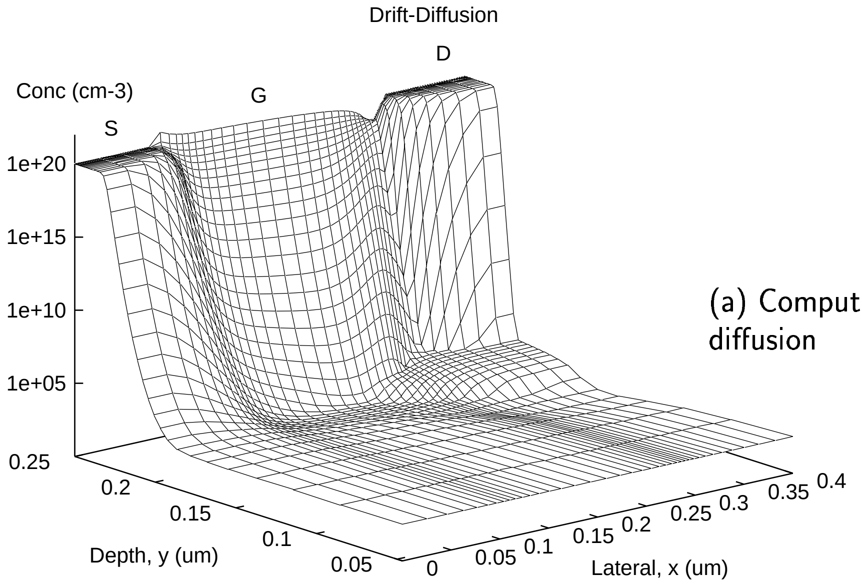
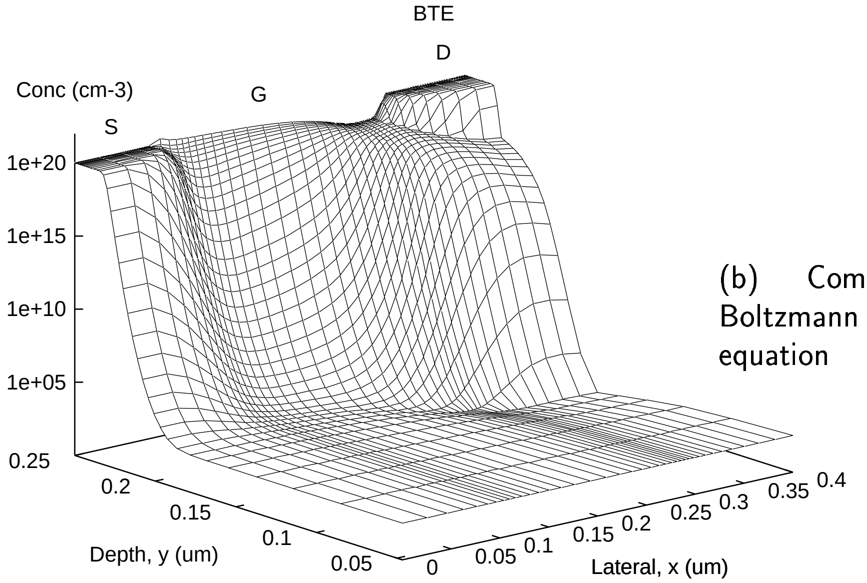


Figure 7.13: Distribution function deep into the MOSFET substrate along a horizontal section at a depth of 0.095μm below the semiconductor-oxide interface. Top figure is for the lower band, band (12); and the lower figure is for the higher bands, band (34). Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.



(a) Computed by drift-diffusion



(b) Computed by Boltzmann transport equation

Figure 7.14: Electron concentration in the MOSFET computed by two methods: (a) Drift-diffusion and (b) Boltzmann transport equation. Electrons density at the drain is more spread in the Boltzmann-equation solution. Gate voltage $V_{GS} =$ 1.5V; drain voltage $V_{DS} =$ 2.5V.

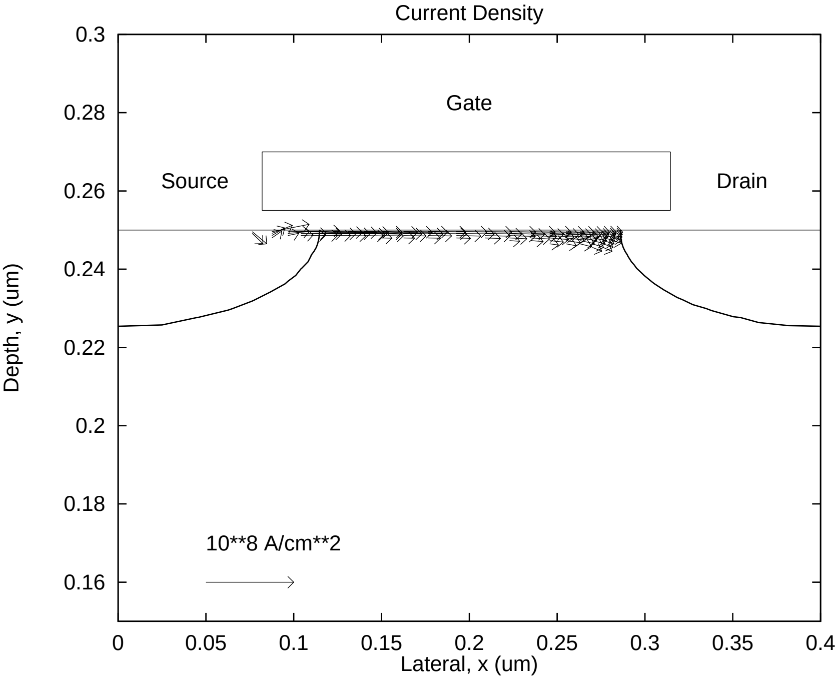
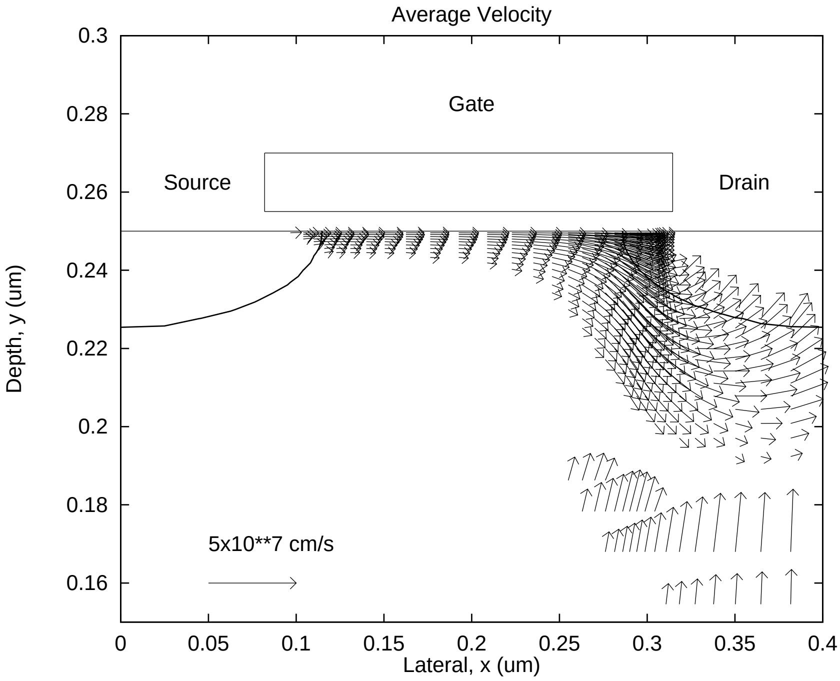
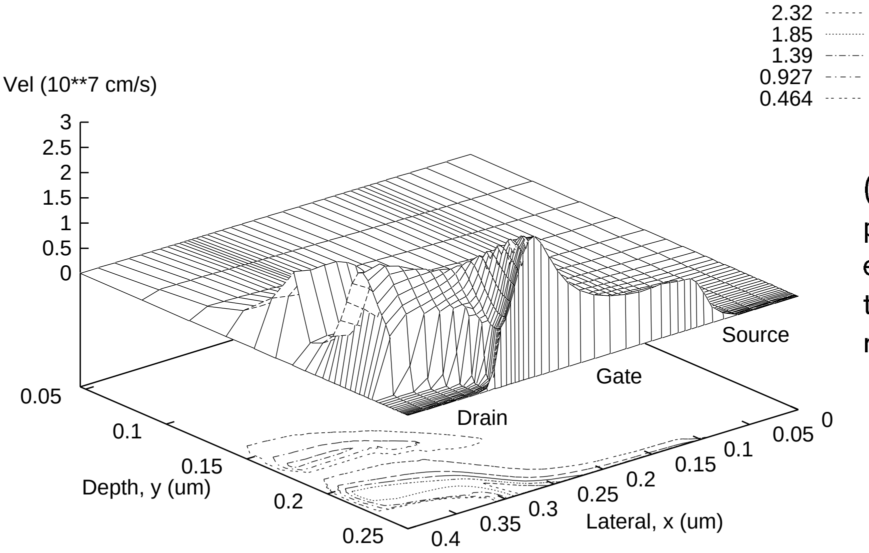


Figure 7.15: Vector plot of the current density, $\mathbf{J}$, in the MOSFET. Current flows only in a thin channel under the oxide. Scale is shown in the lower left corner. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.



(a) Vector plot of average electron velocity



(b) Surface plot of average electron velocity magnitude

Figure 7.16: Electron velocity in the MOSFET. Sub-figure (a) is the vector plot of electron velocity; the scale appears at the lower left corner. Sub-figure (b) is the surface plot of electron velocity magnitude. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

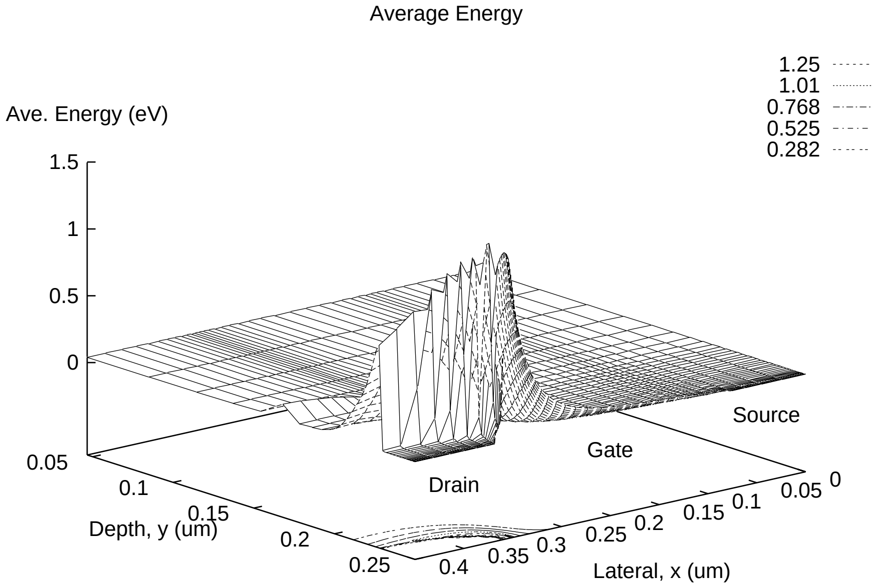


Figure 7.17: Average energy of electrons in the MOSFET. Electrons near the drain have posses large energy; electrons in the substrate and source are at equilibrium thermal energy of $3/2\,k_BT_L$ which, at room temperature of $T_L = 300$K, is 0.0388eV. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

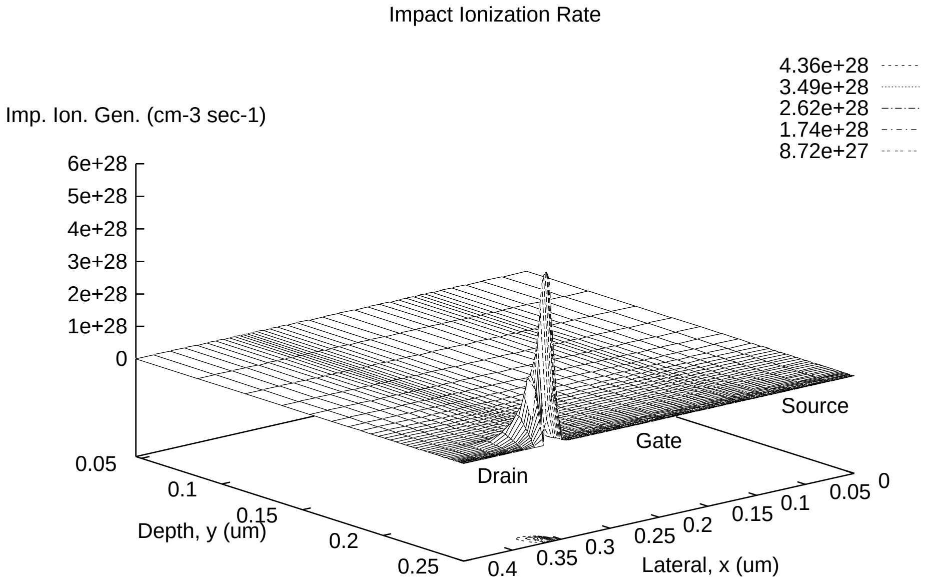


Figure 7.18: Generation rate due to impact ionization in the MOSFET. Most of the generation occurs near the drain. Holes generated by impact ionization cause substrate currents. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

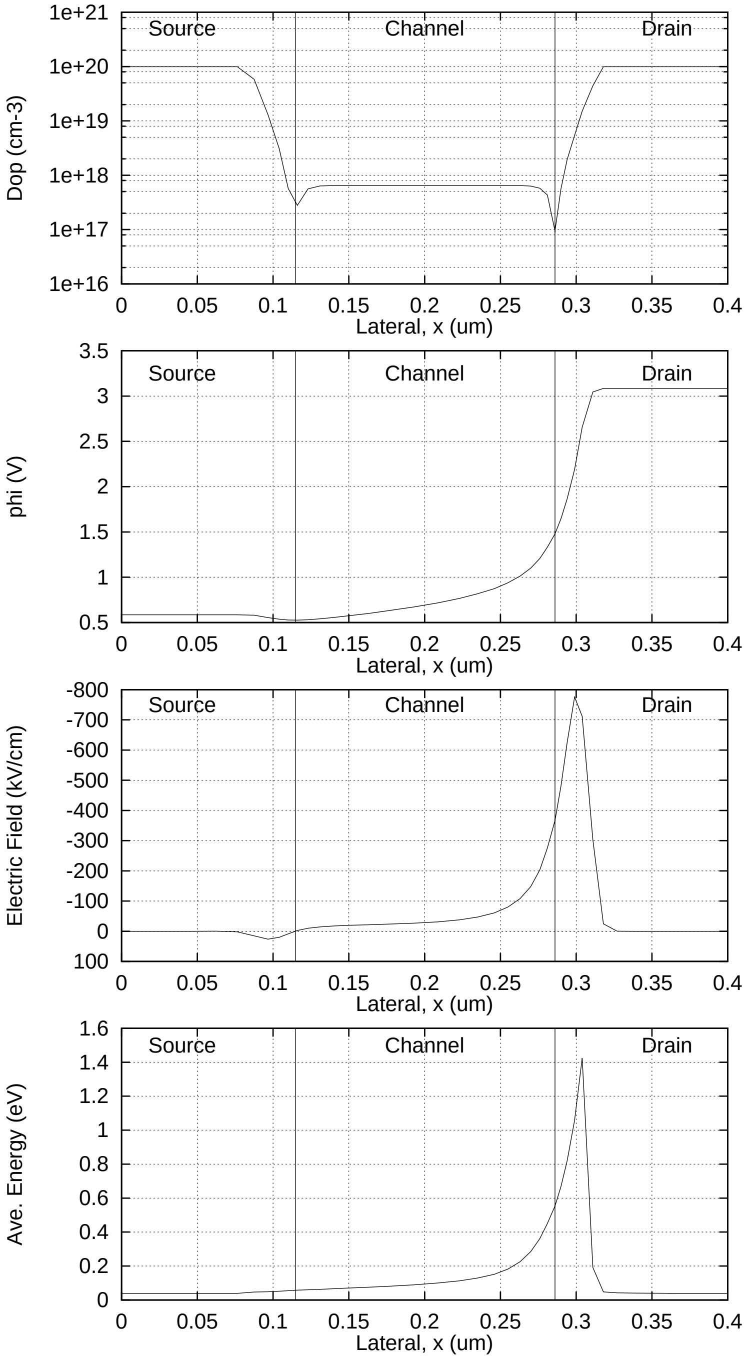


Figure 7.19: Doping profile, electric potential, lateral electric field, average energy under the oxide of the MOSFET. All quantities have been averaged over a depth of 4.96nm. The vertical lines delineate the metallurgical *source*, *channel* and *drain*. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

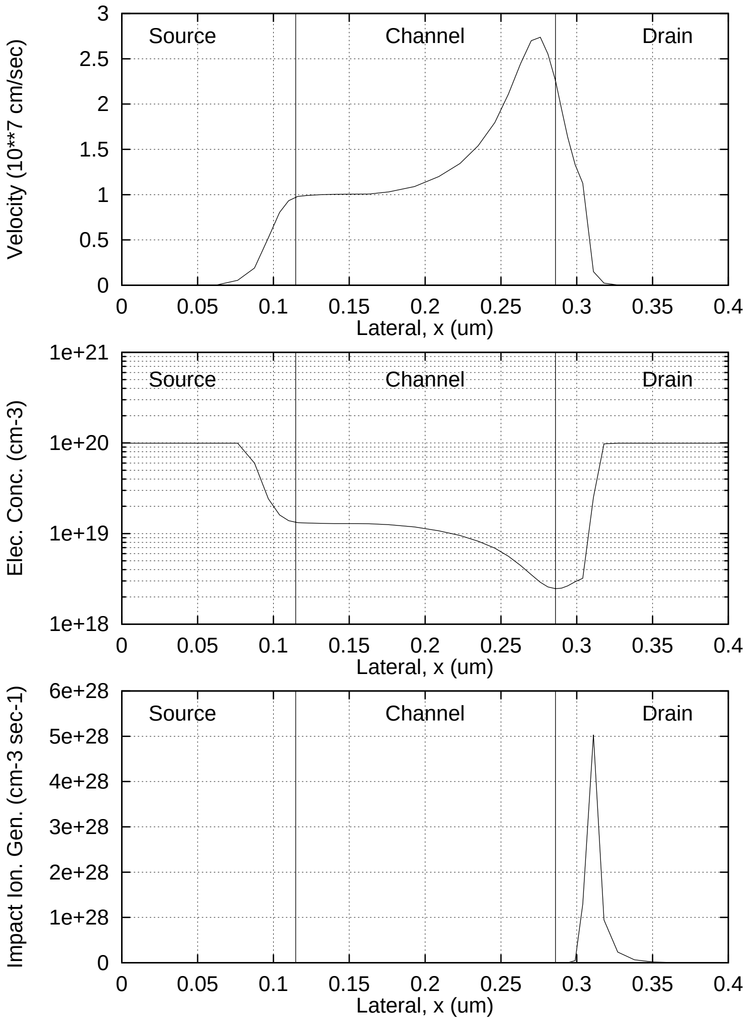


Figure 7.20: Average electron velocity, electron concentration, impact-ionization generation rate under the oxide of the MOSFET. All quantities have been averaged over a depth of 4.96nm. The vertical lines delineate the metallurgical *source*, *channel* and *drain*. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

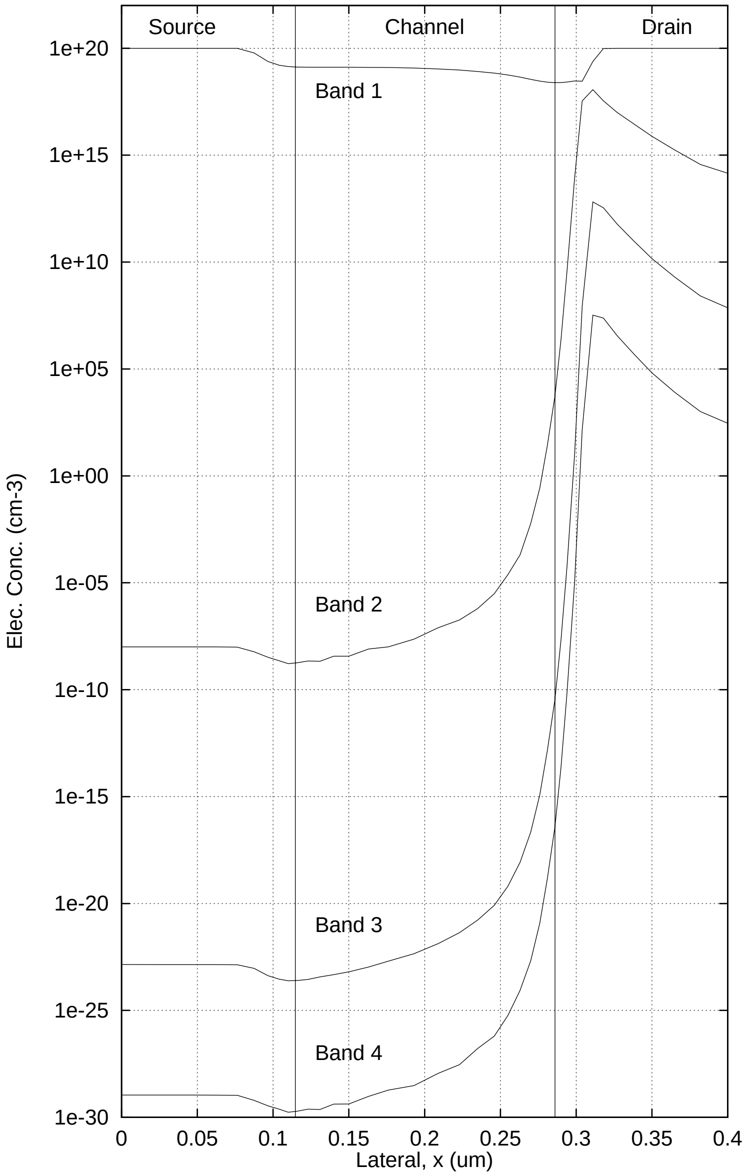


Figure 7.21: Electron concentration in bands (1), (2), (3), and (4) under the oxide of the MOSFET. Electron concentration in the upper bands increases drastically due to both electric field and inter-band scattering. All Concentrations have been averaged over a depth of 4.96nm. The vertical lines delineate the metallurgical *source*, *channel* and *drain*. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

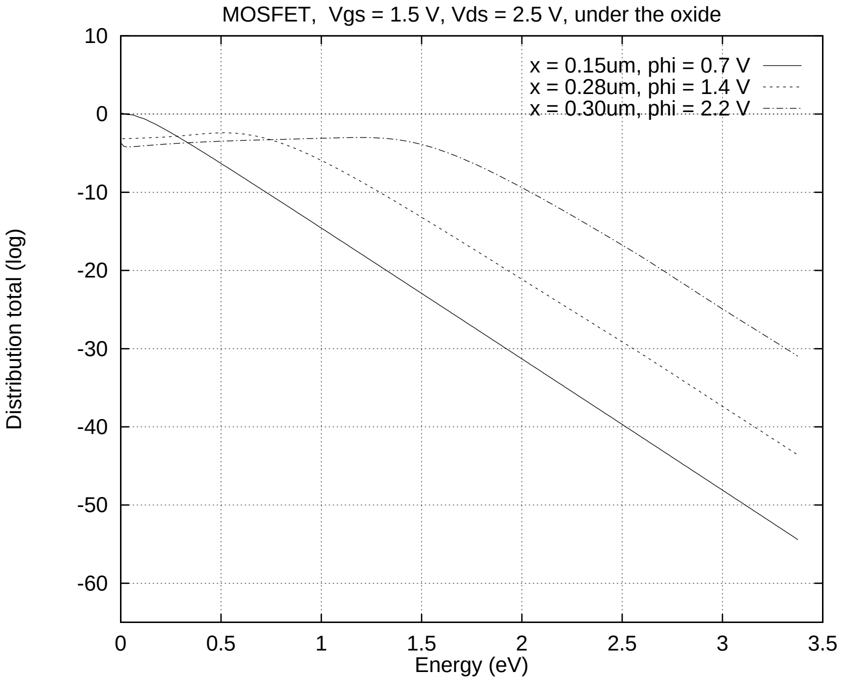


Figure 7.22: Ballistic Transport in the MOSFET. The distribution function is plotted at three points in the channel; the lateral position, x, is 0.15μm, 0.28μm, and 0.3μm; and the depth from the oxide is 4Å for the three points. In fact all three points are on the very first horizontal grid line under the oxide. Electric potential, ϕ, at the three points is 0.7V, 1.4V, and 2.2V respectively. The distribution is a weighted sum of distribution in bands (12) and (34); weighted by density of states. Distributions are shifted from each other in energy by an amount $\Delta\varepsilon = q\,\Delta\phi$, indicating the presence of ballistic transport. Gate voltage $V_{GS} = 1.5$V; drain voltage $V_{DS} = 2.5$V.

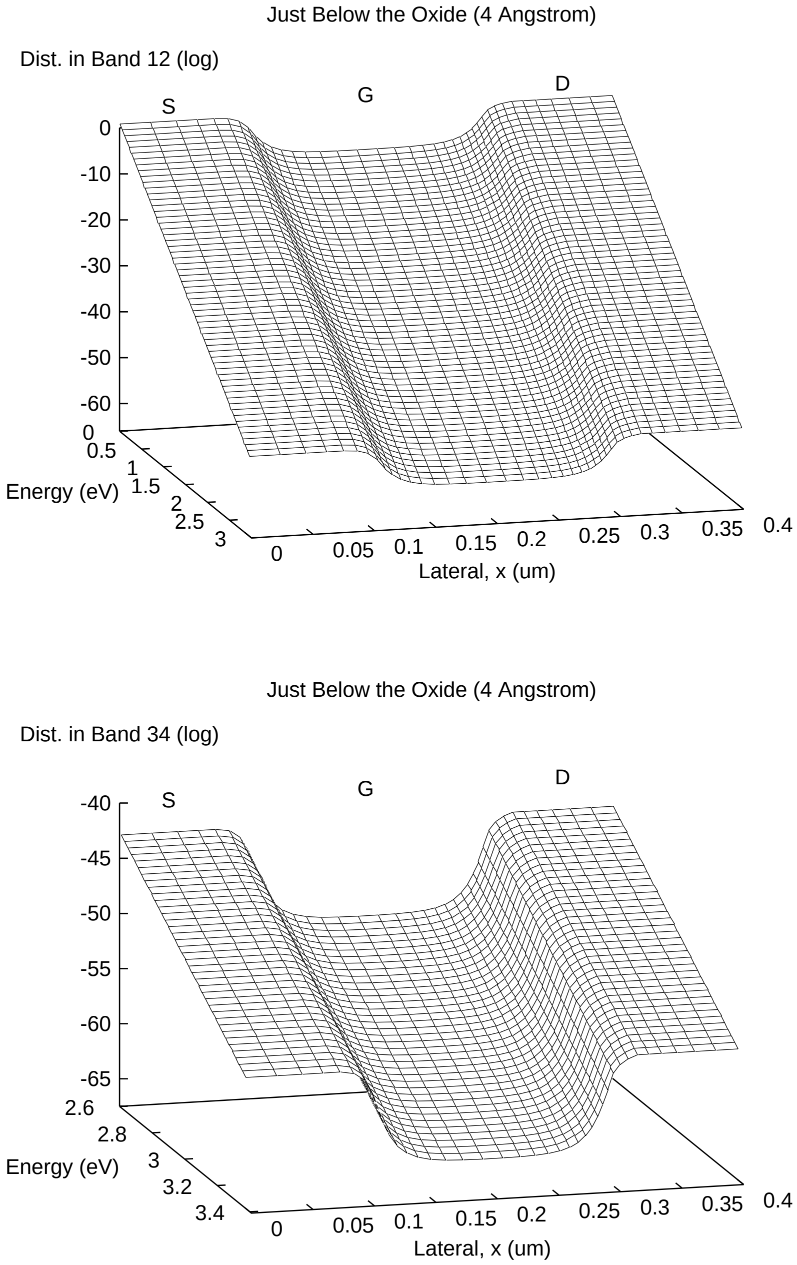
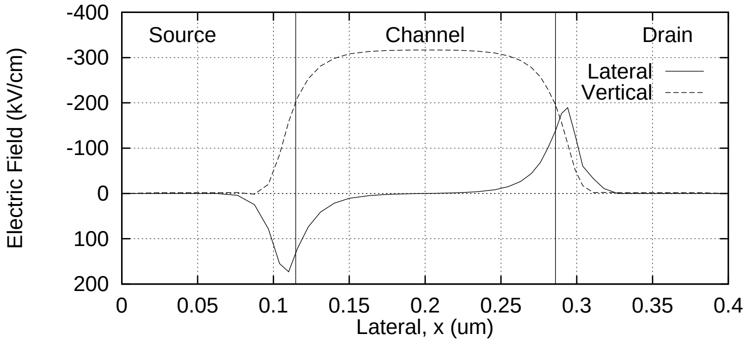
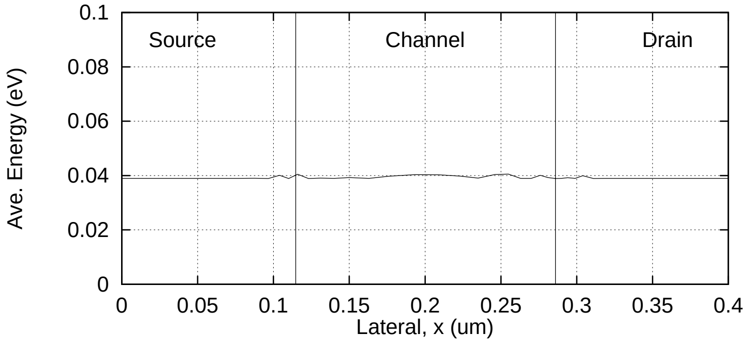


Figure 7.23: Distribution function in the MOSFET at thermal equilibrium. The distribution is shown to be Maxwellian ($\propto \exp(-\varepsilon/k_B T_L)$). Distribution is plotted along a horizontal grid line, 4Å below the oxide. Gate voltage $V_{GS} = 0$V; drain voltage $V_{DS} = 0$V.



(a) Lateral and vertical electric field



(b) Average energy

Figure 7.24: Electric field and average energy at thermal equilibrium in the MOSFET. All quantities are plotted along a horizontal grid line, 4Å below the oxide. Sub-figure (a) plots the lateral and vertical electric fields. Sub-figure (b) plots the average energy, which is virtually constant at the thermal equilibrium value of $3/2\, k_B T_L = 0.0388\text{eV}$ at room temperature. Gate voltage $V_{GS} = 0\text{V}$; drain voltage $V_{DS} = 0\text{V}$.

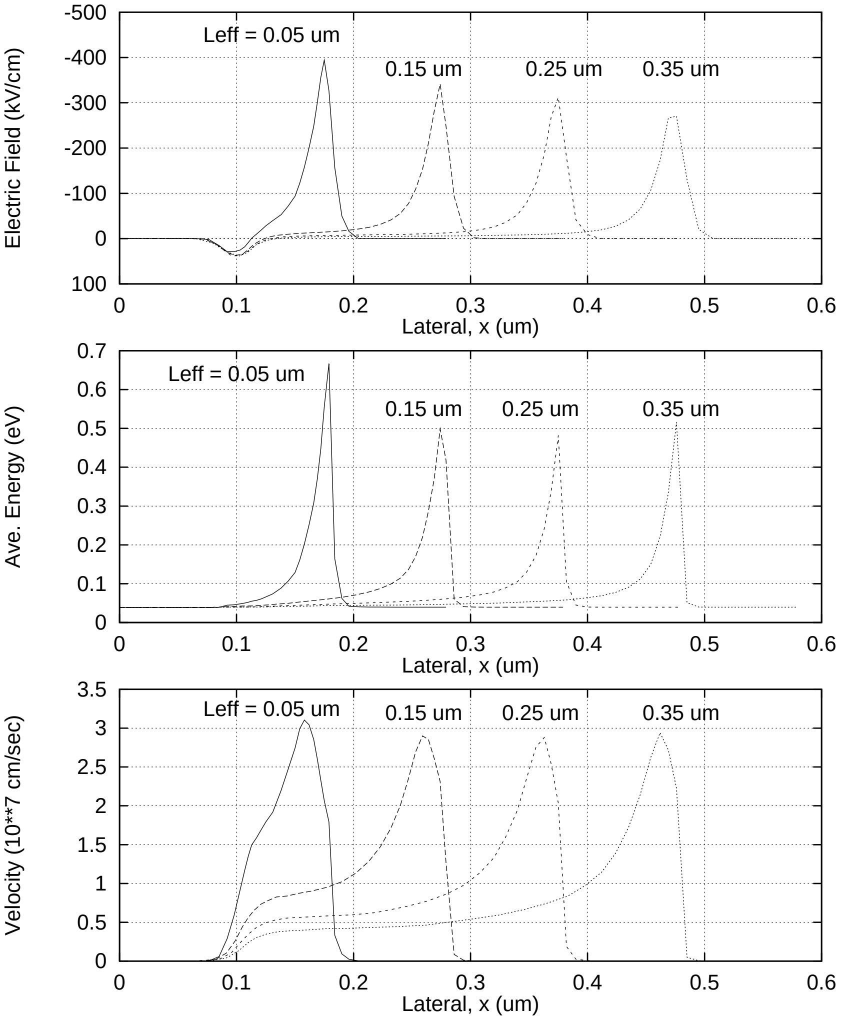


Figure 7.25: Effect of channel-length on MOSFET operation. Three quantities in the MOSFET channel are plotted: Lateral electric field, E_x; average electron energy, $< \varepsilon >$; and average electron velocity. These variables were averaged over a depth of 4.96nm under the oxide. The effective channel-lengths, L_{eff}, which in our case is defined as the metallurgical channel-length, are 0.05μm, 0.15μm, 0.25μm, and 0.35μm. The gate-lengths, L_{gate}, are 0.11μm, 0.21μm, 0.33μm, and 0.42μm rrespectively. Gate voltage $V_{GS} = 1$V; drain voltage $V_{DS} = 1$V for all channel-lengths.

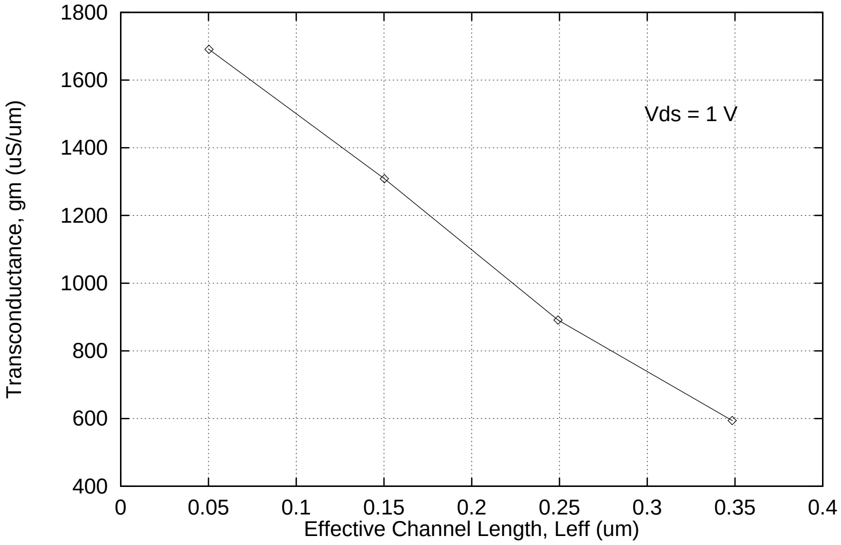


Figure 7.26: Simulated transconductance, g_m, of the MOSFET's as a function of effective channel-length, L_{eff}. Transconductance was calculated by taking the difference of drain currents at V_{GS} =1.0V and 0.8V, with drain voltage V_{DS} = 1.0V for all channel-lengths.

CHAPTER 8

CONCLUSION

In this dissertation the spherical-harmonic-expansion approach to solve the semiconductor Boltzmann transport equation was developed further. The salient features of the dissertation is the use of all four bands from the popular multi-band band-structure of silicon. The Boltzmann equation was solved self-consistently with the Poisson and hole-continuity equation to simulate the one-dimensional n^+nn^+ device and bipolar junction transistor (BJT). The Boltzmann equation was extended to two spatial dimensions to simulate a metal-oxide-semiconductor field-effect transistor (MOSFET). The simulator was applied to investigate non-equilibrium non-local phenomena of velocity overshoot and ballistic transport.

As part of the numerical schemes the most salient contribution is the development of a new diagonally-dominant formulation for the Boltzmann equation. This approach transforms the Boltzmann equation to a new pre-Maxwellian variable. This variable has many desirable qualities: (a) It enhances the Diagonal dominance of the discretized equations. (b) It accounts for the rapid exponential variation of the distribution in both energy and space. (c) In addition, the new pre-Maxwellian variable allows us to write the electron concentrations in Slotboom variables, which opens the possibility of using superior Poisson equation solvers like Mayergoyz' fixed-point algorithm. (d) The transformed equations retain their linearity.

It was found that the commonly-used elastic approximation for the acoustic phonons resulted in spurious noise in the computed distribution function. This numerical noise was analyzed and corrected by incorporating acoustic phonons in inelastic approximation. To this end, a new formulation for inelastic acoustic phonons was developed. This formulation is self-adjoint, and is ideally suited for numerical simulation.

The multi-band Boltzmann equations were discretized by a current-conserving control-volume discretization. This discretization preserves the conservative nature of the spherical-harmonic Boltzmann equations. The discretized multi-band Boltzmann equations were solved by a line successive-over-relaxation (SOR) method. The coupled system was solved by the Gummel's decoupled method.

In addition, the interesting case of thermal equilibrium was discussed. It was proven that the analytical solution of the multi-band Boltzmann equations is Maxwellian distribution. The thermal-equilibrium was used to test the code, as well as studied as an interesting phenomena in its own right.

APPENDIX A

CURRENT-CONSERVATION PROPERTY OF THE SPHERICAL-HARMONIC BOLTZMANN EQUATION

A.1 Proof of Current-Continuity

The Boltzmann equation, after a first-order spherical-harmonic expansion of the distribution function, reads (Section 2.3, Eq. (2.24)):

$$\frac{1}{3g}\frac{\partial}{\partial x}\left(\tau u_g^2 g \frac{\partial F_0^0}{\partial x}\right) + \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = 0 \tag{A.1}$$

The aim of this appendix is to prove that Eq. (A.1) satisfies the electron current-continuity equation:

$$\frac{\partial J(x)}{\partial x} = G(x) = G_{\text{impact}}(x) \tag{A.2}$$

where $J(x)$ is the current density at x, measured as number of electrons per unit area per unit time; and $G(x)$ is the net generation of electrons at x, measured as electrons per unit volume per unit time. The only generation mechanism we consider in this dissertation is impact-ionization, therefore G has been set equal to G_{impact}.

To prove current-continuity property, we multiply Eq. (A.1) by $3g$ and integrate from Hamiltonian $H_{\text{min}} = -q\phi(x)$ to $H_{\text{max}} = \varepsilon_{\text{max}} - q\phi(x)$

$$\int_H \frac{\partial}{\partial x}\left(\kappa \frac{\partial F_0^0}{\partial x}\right) dH + \int_H 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} dH = 0 \tag{A.3}$$

where we have set $\kappa = \tau u_g^2 g$. Current density is known from Eq. (2.63):

$$J(x) = -c_j \int_H \kappa \frac{\partial F_0^0}{\partial x}\, dH \tag{A.4}$$

where $c_j = 1/(24\pi^3\sqrt{\pi})$. Multiplying Eq. (A.3) by $-c_j$ and using Eq. (A.4):

$$-c_j \frac{\partial}{\partial x} \int_H \kappa \frac{\partial F_0^0}{\partial x}\, dH - c_j \int_H 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} dH = 0 \tag{A.5}$$

$$\frac{\partial J(x)}{\partial x} - c_j \int_H 3g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} dH = 0 \tag{A.6}$$

The collision term for the scattering mechanism is:

$$\left[\frac{\partial F_0^0}{\partial t}\right]_{\text{collision}} = \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{opt}} + \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} + \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{impurity}} + \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{impact}} \tag{A.7}$$

Scattering due to inter-band optical phonons is not written because we are implicitly considering only one conduction band; including it is not important, it has the same form as optical phonon term, therefore same conclusions would apply to it (refer to item 4 on the facing page). Ionized impurity scattering term, Eq. (2.83), is identically equal to zero, therefore it need not be considered any further.

It is easy to prove the following identities:

$$\int_H g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{opt}} dH \equiv 0 \tag{A.8}$$

$$\int_H g \left[\frac{\partial F_0^0}{\partial t}\right]_{\text{ac}} dH \equiv 0 \tag{A.9}$$

This mathematical identity is physically sound: It tells us that that there is no *net* generation due to either optical or acoustic phonons, The only non-zero term left in the collision terms is impact ionization scattering. This, by considering Eq. (2.96) and Eq. (2.97), can be written as follows:

$$\begin{aligned} c_j \int_H 3g \left[\frac{\partial F_0^0(H)}{\partial t}\right]^{\text{total}}_{\text{impact}} dH &= \frac{3}{24\pi^3\sqrt{\pi}} \int_H g \left[\frac{\partial F_0^0(H)}{\partial t}\right]^{\text{out}}_{\text{impact}} dH \\ &= \int_H \tilde{h}\, \frac{F_0^0(H)}{\tau_{\text{impact}}(H+q\phi)}\, dH \\ &= G_{\text{impact}} \end{aligned} \tag{A.10}$$

where we used the relation $\tilde{h}(\varepsilon) = (1/\sqrt{4\pi})(1/4\pi^3)g(\varepsilon)$ from Section 2.5.5. We can now substitute Eq. (A.10) in Eq. (A.6) and see that it is same as the current-continuity condition, Eq. (A.2), we set out to prove. This concludes the proof.

A.2 Discussion

Following points are worth noting:

1. The original Boltzmann equation, Eq. (2.5), is a statement of conservation of electrons: it tracks particles in k-space as well as real-space. The spherical-harmonic Boltzmann equation, Eq. (2.24) or Eq. (A.1), is also a conservation equation: it tracks particles in energy or Hamiltonian as well as real-space. The conserving nature of the equation is not lost by spherical-harmonic expansion.
2. It would be desirable to have a discretization which also ensures this current continuity. The control-volume discretization of Chapter 4 is ideally suited.
3. Identities Eq. (A.8) and Eq. (A.9) indicate that there is no *net* generation of electrons due to optical or acoustic phonons. Generation of electrons in one region of energy is due to loss of electrons from other regions. The total net loss or generation is zero.
4. Inter-band phonons have the same functional form as optical phonons. Total net generation due to inter-band phonons is also zero. If we consider inter-band phonons we could write an identity similar to Eq. (A.8):

   $$\sum_{\nu} \int_{H} g^{(\nu)} \left[\frac{\partial F_0^{0\,(\nu)}}{\partial t} \right]_{\text{ib}} dH \equiv 0 \tag{A.11}$$

   where (ν) is the band index.
5. Identity Eq. (A.8) has a physical interpretation. It is saying that the in a energy or Hamiltonian range ΔH the net generation of electrons due to optical phonons is proportional to $g \left[\partial F_0^0 / \partial t\right]_{\text{opt}} \Delta H$.
6. Using the identity from Eq. (A.9) the same may be said about acoustic phonons: net generation of electrons due to acoustic phonons is proportional to $g \left[\partial F_0^0 / \partial t\right]_{\text{ac}} \Delta H$.
7. Satisfying the identities Eq. (A.8), Eq. (A.9), and Eq. (A.11) is crucial to proving current-conservation. A numerical scheme which aspires for current-continuity must also satisfy these identities.

APPENDIX B

MULTI-BAND BOLTZMANN EQUATION AT THERMAL EQUILIBRIUM

In this appendix the analytical solution of the Boltzmann-Poisson system at thermal equilibrium is derived. At thermal equilibrium, we will prove, the analytical solution of the multi-band Boltzmann equation is Maxwellian distribution function; and the electric potential satisfies the equilibrium non-linear Poisson equation.

The proof starts by assuming that the distribution at thermal equilibrium is Maxwellian, which is shown to satisfy the Boltzmann transport and non-linear Poisson equation. Since the Boltzmann-Poisson system has a unique solution, this is the only solution at thermal equilibrium.

Thermal equilibrium is defined as a situation when there are no currents in the device: $\mathbf{J}(\mathbf{r}) \equiv 0$. The collision integral for a non-degenerate semiconductor can be written as (same as Eq. (2.65)):

$$\left[\frac{\partial f(\mathbf{k})}{\partial t}\right]_{\text{collision}} = \frac{\Omega}{(2\pi)^3}\int_{\mathbf{k}'} f(\mathbf{k}')S(\mathbf{k}',\mathbf{k}) - f(\mathbf{k})S(\mathbf{k},\mathbf{k}')\; d^3\mathbf{k}' \tag{B.1}$$

Here $S(\mathbf{k},\mathbf{k}')$ is the probability of an electron scattering from state $\mathbf{k}$ to state $\mathbf{k}'$, and it satisfies the following relationship [73, 77]:

$$S(\mathbf{k},\mathbf{k}')\exp\left(-\frac{\varepsilon^{(\nu)}(\mathbf{k})}{k_B T_L}\right) = S(\mathbf{k}',\mathbf{k})\exp\left(-\frac{\varepsilon^{(\nu')}(\mathbf{k}')}{k_B T_L}\right) \tag{B.2}$$

where $\varepsilon^{(\nu)}(\mathbf{k})$ represents the ε-$\mathbf{k}$ relationship for band (ν). Let us assume that the distribution function in two bands, bands (12) and (34) is given by

$$f^{(12)}(\mathbf{r},\mathbf{k}) = A^{(12)} n^{(12)}(\mathbf{r})\exp\left(-\frac{\varepsilon^{(12)}(\mathbf{k})}{k_B T_L}\right) \tag{B.3}$$

$$f^{(34)}(\mathbf{r},\mathbf{k}) = A^{(34)} n^{(34)}(\mathbf{r}) \exp\left(-\frac{\varepsilon^{(34)}(\mathbf{k})}{k_B T_L}\right) \tag{B.4}$$

We choose the constants A such that $1/4\pi^3 \int f(\mathbf{k})\, d^3\mathbf{k} = n$, which leads to

$$\frac{1}{A^{(12)}} = \frac{1}{4\pi^3} \int_{\text{Band 12}} \exp\left(-\frac{\varepsilon^{(12)}(\mathbf{k})}{k_B T_L}\right) d^3\mathbf{k} \tag{B.5}$$

$$\frac{1}{A^{(34)}} = \frac{1}{4\pi^3} \int_{\text{Band 34}} \exp\left(-\frac{\varepsilon^{(34)}(\mathbf{k})}{k_B T_L}\right) d^3\mathbf{k} \tag{B.6}$$

If we substitute the distribution of Eq. (B.3) and Eq. (B.4) in the Boltzmann transport equation Eq. (2.5), the collision term Eq. (B.1) is nullified due to Eq. (B.2) and an additional constraint of $A^{(12)} n^{(12)} = A^{(34)} n^{(34)}$. For band (12) this results in

$$\frac{1}{\hbar}\nabla_k \varepsilon \cdot \nabla_r \left\{ A_{(12)} n^{(12)}(\mathbf{r}) \exp\left(-\frac{\varepsilon^{(12)}(\mathbf{k})}{k_B T_L}\right)\right\} - \frac{q\mathbf{E}(\mathbf{r})}{\hbar} \cdot \nabla_k \left\{ \exp\left(-\frac{\varepsilon^{(12)}(\mathbf{k})}{k_B T_L}\right)\right\} = 0 \tag{B.7}$$

Using $\mathbf{E}(\mathbf{r}) = -\nabla_r \phi(\mathbf{r})$ we can analytically solve the above equation, which gives

$$n^{(12)}(\mathbf{r}) = n_i^{(12)} \exp\left(\frac{\phi(\mathbf{r})}{V_t}\right) \tag{B.8}$$

where $V_t = k_B T_L / q$ is the thermal voltage, and $n_i^{(12)}$ is a constant. Similarly for band (34) we get:

$$n^{(34)}(\mathbf{r}) = n_i^{(34)} \exp\left(\frac{\phi(\mathbf{r})}{V_t}\right) \tag{B.9}$$

By substituting the total electron concentration $n(\mathbf{r}) = n^{(12)}(\mathbf{r}) + n^{(34)}(\mathbf{r})$ in the Poisson equation we obtain:

$$\nabla_r^2 \phi(\mathbf{r}) = \frac{q}{\epsilon_{\text{si}}} \left\{ \left(n_i^{(12)} + n_i^{(34)}\right) \exp\left(\frac{\phi(\mathbf{r})}{V_t}\right) - p(\mathbf{r}) - \mathcal{D}(\mathbf{r}) \right\} \tag{B.10}$$

This, as we can see, is the non-linear equilibrium Poisson equation. The sum of constants $n_i^{(12)} + n_i^{(34)}$ can be interpreted as the intrinsic carrier density n_i.

The above exercise has shown that if the distribution function is of the Maxwellian form as given by Eq. (B.3) and Eq. (B.4), then it satisfies the Boltzmann equation as well as the the non-linear equilibrium Poisson equation. Since the coupled Boltzmann-Poisson system can be shown to possess a unique solution [77], this is the only solution at thermal equilibrium. Thus, for the thermal equilibrium case, we have obtained an analytical solution of the multi-band Boltzmann equation. The resultant distribution function is Maxwellian, given by Eq. (B.3) and Eq. (B.4).

APPENDIX C

HARMONIC-MEAN SCHEME

In this appendix we derive the harmonic-mean scheme of Patankar [113,116]. We will also point out the similarity of this harmonic-mean scheme to the Scharfetter-Gummel discretization scheme and another scheme used for Boltzmann equation, the Liang-Goldsman-Mayergoyz scheme.

C.1 Derivation

Let us consider the one-dimensional equation

$$\frac{\partial J}{\partial x} = S \tag{C.1}$$

where S is the source term; and current, or the flux, J, is given by

$$J = -\kappa \frac{\partial u}{\partial x} \tag{C.2}$$

where u is a state variable, and κ the "conductivity." Eq. (C.1) and Eq. (C.2) could be thought of as a generalized representation for many equations: (a) the spherical-harmonic Boltzmann equation, Eq. (2.24) or Eq. (4.27); (b) heat equation, Eq. (4.28), [113]; or (c) drift-diffusion equations in Slotboom variables [98,102,104,105].

Let us consider the grid shown in Fig. C.1 on the next page. Integrating Eq. (C.1) from x_w to x_e

$$\int_e^w \frac{\partial J}{\partial x}\, dx = \int_e^w S\, dx \tag{C.3}$$

leads to

$$\frac{J_e - J_w}{\Delta x_P} = \bar{S} \Delta x_P \tag{C.4}$$

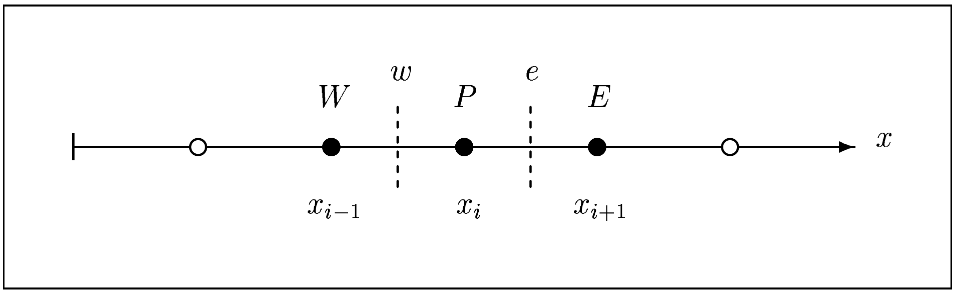


Figure C.1: Grid for derivation of the harmonic-mean scheme. Filled circles indicate the computational molecule.

where $\Delta x_P = x_e - x_w$, and $\bar{S}$ is the average of the source term in the interval. J_e and J_w are the fluxes at the interface e and w respectively. The problem is now to approximate these fluxes.

By rearranging Eq. (C.2) and integrating from points P to E, we get

$$\kappa^{-1} J = -\frac{\partial u}{\partial x} \tag{C.5}$$

$$\int_P^E \kappa^{-1} J \, dx = -\int_P^E \frac{\partial u}{\partial x} \, dx = -\left(u_E - u_P\right) \tag{C.6}$$

Now we make critical assumption: flux J is constant between point P and E. This allows us to take J out of the integral:

$$J \int_P^E \kappa^{-1} \, dx = -\left(u_E - u_P\right) \tag{C.7}$$

which leads to

$$J = -\frac{u_E - u_P}{\displaystyle\int_P^E \kappa^{-1} \, dx} \tag{C.8}$$

Now it boils down to assuming a profile for κ [113]. If we assume that $\kappa =$ constant $= (\kappa_P + \kappa_E)/2$ in the interval from P to E, then we obtain the normal *arithmetic-mean scheme*:

$$J = -\frac{\kappa_P + \kappa_E}{2} \frac{u_E - u_P}{\Delta x_e} \tag{C.9}$$

where $x_e = x_E - x_P$.

If, however, we make the assumption that κ is constant in the control-volume surrounding the grid points, that is $\kappa = \kappa_P$ for $x \in [x_w, x_e]$, we get

$$\begin{aligned} \int_P^E \kappa^{-1}\, dx &= \int_P^e \kappa_P^{-1}\, dx + \int_e^E \kappa_E^{-1}\, dx \\ &= \kappa_P^{-1}\, \frac{\Delta x_e}{2} + \kappa_E^{-1}\, \frac{\Delta x_e}{2} \\ &= \frac{\kappa_E + \kappa_E}{2\, \kappa_P\, \kappa_E}\, \Delta x_e \end{aligned} \tag{C.10}$$

Substituting Eq. (C.10) into Eq. (C.8) leads to the *harmonic-mean scheme*:

$$J = -\, \frac{2\, \kappa_P\, \kappa_E}{\kappa_P + \kappa_E}\, \frac{u_E - u_P}{\Delta x_e} \tag{C.11}$$

The interface conductivity has been approximated by a harmonic mean;

$$\boxed{\kappa_e = \frac{2\, \kappa_P\, \kappa_E}{\kappa_P + \kappa_E}} \tag{C.12}$$

This scheme has many desirable properties, as discussed in [113,116]. It is ideal to use in situations with rapid variations in conductivity κ. Following subsections point out the similarity of this scheme to other discretization schemes.

C.2 Similarity to Scharfetter-Gummel Discretization

An astute reader may notice the similarity between the harmonic-mean scheme and the Scharfetter-Gummel discretization [100]. In the derivation for both schemes, we make the same crucial assumption: flux J is constant between the grid points. After making this assumption, we solve the in-homogeneous ordinary differential equation to obtain a profile of the dependent variable (refer to Section E.2 on page 190 for an application to the hole-continuity equation). The only difference is that the starting equation for Scharfetter-Gummel discretization, in addition to having the diffusion term, also has the convection term. The harmonic-mean scheme and the Scharfetter-Gummel, therefore, arise from the *same* basic assumption [106]. [1]

---

[1] The similarity between harmonic-mean and the Scharfetter-Gummel scheme was pointed out to me by Dr. Paul M. de Zeeuw of CWI, Netherlands.

C.3 Similarity to Liang-Goldsman-Mayergoyz Scheme

If, instead of assuming a constant, or piece-wise constant, profile of κ in Eq. (C.8), we assume a *linear* profile, we obtain the scheme used by Liang, Goldsman, and Mayergoyz for the Boltzmann equation [43, 51].

APPENDIX D

POISSON EQUATION

D.1 The Boundary-Value Problem

From the usual electromagnetic theory, the divergence of displacement vector is equal to the charge density [72]:

$$\nabla_r \cdot \mathbf{D} = \nabla_r \cdot (\epsilon_{\text{si}}\mathbf{E}) = \rho \tag{D.1}$$

where ϵ_{si} is permittivity of silicon, and $\epsilon_{\text{si}} = \epsilon_{\text{rel}}\epsilon_0$. Relative permittivity of silicon $\epsilon_{\text{rel}} = 11.7$, and permittivity of free space $\epsilon_0 = 8.85418 \times 10^{-12}$ F/m. ρ is the total charge density at $\mathbf{r}$, and is given by:

$$\rho(\mathbf{r}) = q\left(p(\mathbf{r}) - n(\mathbf{r}) + N_D^+(\mathbf{r}) - N_A^-(\mathbf{r})\right) \tag{D.2}$$

Where $n(\mathbf{r})$ and $p(\mathbf{r})$ are the electron and hole concentration at $\mathbf{r}$. $N_D^+(\mathbf{r})$ and $N_A^-(\mathbf{r})$ are ionized impurity concentration for donor and acceptor type. We can assume from now on that all the impurities are ionized, hence we will drop the superscripts $^+$ and $^-$. Furthermore, since it is the difference of these dopings that is of importance, we can define a total net doping $\mathcal{D}(\mathbf{r}) \triangleq N_D(\mathbf{r}) - N_A(\mathbf{r})$. Expressing the electric field as the gradient of a scalar electric potential, $\phi(\mathbf{r})$,

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= -\nabla_r\phi(\mathbf{r}) \\ &= -\frac{\partial\phi}{\partial x}\,\hat{\mathbf{i}} - \frac{\partial\phi}{\partial y}\,\hat{\mathbf{j}} - \frac{\partial\phi}{\partial z}\,\hat{\mathbf{k}}\end{aligned} \tag{D.3}$$

we get the linear Poisson equation for semiconductors:

$$\boxed{\nabla_r^2\phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_{\text{si}}} = \frac{q}{\epsilon_{\text{si}}}\left(n(\mathbf{r}) - p(\mathbf{r}) - \mathcal{D}(\mathbf{r})\right)} \tag{D.4}$$

As an aside we may also write down the form of the non-linear Poisson equation [95, 98]. If we express the carrier concentrations in terms of the Slotboom variables u and v,

$$n(\mathbf{r}) = n_i\, u\, e^{\phi(\mathbf{r})/V_t} \qquad p(\mathbf{r}) = n_i\, v\, e^{-\phi(\mathbf{r})/V_t} \tag{D.5}$$

and substitute in Eq. (D.4), we get the non-linear Poisson equation:

$$\nabla_r^2 \phi(\mathbf{r}) = \frac{q}{\epsilon_{\text{si}}} \left(n_i\, u\, e^{\phi(\mathbf{r})/V_t} - n_i\, v\, e^{-\phi(\mathbf{r})/V_t} - \mathcal{D}(\mathbf{r}) \right) \tag{D.6}$$

It has been reported that the non-linear Poisson equation in self-consistent Monte Carlo simulation results in faster convergence [70]. This is perhaps because the non-linear Poisson equation anticipates the electron-concentration change with the changing electric potential [52]. In this dissertation, however, we have only used the linear Poisson equation; rest of this Appendix focuses on that.

D.1.1 Boundary Conditions

An ohmic contact to the semiconductor is the most the most common boundary condition that arises for the Poisson equation. This boundary condition takes on a Dirichlet form:

$$\begin{aligned} \phi &= V_{\text{applied}} + \phi_{\text{built in}} \\ &= V_{\text{applied}} + \frac{k_B T_L}{q} \frac{\mathcal{D}}{|\mathcal{D}|} \ln \left(\frac{|\mathcal{D}|}{2n_i} + \sqrt{\frac{\mathcal{D}^2}{4n_i^2} + 1} \right) \end{aligned} \tag{D.7}$$

where n_i is the intrinsic concentration.

We may also have a boundary condition where the electric field is specified, usually zero. The boundary condition takes the Neumann form; zero Neumann if the right-side is zero:

$$\nabla_r \phi = -E_{\text{specified}} \tag{D.8}$$

D.2 Discretization

One-dimensional Poisson equation reads:

$$\frac{d^2\phi(x)}{dx^2} = \frac{q}{\epsilon_{\text{si}}} \left(n(x) - p(x) - \mathcal{D}(x) \right) \tag{D.9}$$

We construct a control-volume grid in Fig. D.1 on the next page. Integrating in the control-volume from x_w to x_e

$$\int_w^e \frac{d^2\phi(x)}{dx^2}\, dx = \int_w^e \frac{q}{\epsilon_{\text{si}}} \left(n(x) - p(x) - \mathcal{D}(x) \right)\, dx \tag{D.10}$$

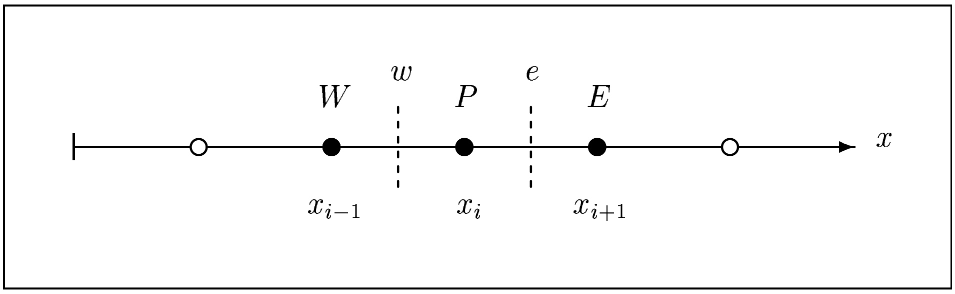


Figure D.1: Grid for the discretization of the Poisson equation. Filled circles indicate the computational molecule.

The left-hand side can be integrated:

$$\int_w^e \frac{d^2\phi(x)}{dx^2}\,dx = \left(\frac{d\phi(x)}{dx}\right)_e - \left(\frac{d\phi(x)}{dx}\right)_w \tag{D.11}$$

The first-order derivative can be approximated by a central-difference discretization. This discretization scheme can be easily shown to be is second-order accurate [114]:

$$\left(\frac{d\phi(x)}{dx}\right)_e = \frac{\phi_E - \phi_P}{\Delta x_e} + \mathcal{O}\left((\Delta x_e)^2\right) \tag{D.12}$$

$$\left(\frac{d\phi(x)}{dx}\right)_w = \frac{\phi_P - \phi_W}{\Delta x_w} + \mathcal{O}\left((\Delta x_w)^2\right) \tag{D.13}$$

Where $\Delta x_e = x_E - x_P$, and $\Delta x_w = x_P - x_W$.

To approximate the integral on the right-hand side of the Poisson equation, Eq. (D.10), we can assume that the average value of the variables of n, p, and $\mathcal{D}$ in the control-volume around the point P is the same as the nodal value of the variables at point P. Such an approximation is second-order accurate. To illustrate this, let us consider a quantity $S(x)$, then

$$\int_w^e S\,dx = \bar{S}\Delta x_P = S_P\Delta x_P + \mathcal{O}\left((\Delta x_P)^2\right) \tag{D.14}$$

Where $\Delta x_P = x_e - x_w = (x_E - x_W)/2$, and $S_P = S(x_P)$ This approximation is second-order accurate. The right side of the Poisson equation, Eq. (D.10), is then approximated as

$$\int_w^e \frac{q}{\epsilon_{\text{si}}}\left(n(x) - p(x) - \mathcal{D}(x)\right)\,dx = \frac{q}{\epsilon_{\text{si}}}\left(n_P - p_P - \mathcal{D}_P\right)\Delta x_P + \mathcal{O}\left((\Delta x_P)^2\right) \tag{D.15}$$

Where the subscript P indicates that the n_P, p_P, and $\mathcal{D}_P$ are evaluated at the point x_P.

Putting it together we have the following second-order-accurate discretization of the Poisson equation:

$$\frac{1}{\Delta x_P}\left(\frac{\phi_E-\phi_P}{\Delta x_e}-\frac{\phi_P-\phi_W}{\Delta x_w}\right)=\frac{q}{\epsilon_{\text{si}}}\left(n_P-p_P-\mathcal{D}_P\right) \tag{D.16}$$

$$\frac{1}{\Delta x_P}\left(\frac{\phi_W}{\Delta x_w}-\left(\frac{1}{\Delta x_w}+\frac{1}{\Delta x_e}\right)\phi_P+\frac{\phi_E}{\Delta x_e}\right)=\frac{q}{\epsilon_{\text{si}}}\left(n_P-p_P-\mathcal{D}_P\right) \tag{D.17}$$

D.2.1 Damping Term

There are serious numerical oscillations when we solve the Boltzmann and Poisson equation self-consistently [33, 48, 70]. The reason for this is easy to understand: potential ϕ depends on the electron concentration, n; electron concentration varies roughly exponentially with potential, $n \sim \exp(\phi/kt)$. This coupling is very strong and the solution oscillates violently, and soon diverges. The damping technique from (refer to p. 211 of [72]) was used in the self-consistent spherical-harmonic simulations by [33, 42, 43, 48, 51].

If we are solving the Poisson equation iteratively, like in a Gummel loop, we can use the solution at the previous iteration to damp the solution in the current iteration. Damping term we add is $r_P^{(m)}(\phi_P^{(m)}-\phi_P^{(m-1)})$, where the damping coefficient $r_P^{(m)}$ at the m^{th} iteration is given by

$$r_P^{(m)}=\frac{n_P^{(m-1)}+p_P^{(m-1)}}{V_t}\cdot\frac{q}{\epsilon_{\text{si}}} \tag{D.18}$$

where $V_t = k_B T_L/q$, the thermal voltage. This damping is added to the Poisson equation

$$\frac{1}{\Delta x_P}\left(\frac{\phi_W^{(m)}}{\Delta x_w}-\left(\frac{1}{\Delta x_w}+\frac{1}{\Delta x_e}\right)\phi_P^{(m)}+\frac{\phi_E^{(m)}}{\Delta x_e}\right)-\overbrace{r_P^{(m)}\left(\phi_P^{(m)}-\phi_P^{(m-1)}\right)}^{\text{Damping term}}$$
$$=\frac{q}{\epsilon_{\text{si}}}\left(n_P^{(m)}-p_P^{(m)}-\mathcal{D}_P\right) \tag{D.19}$$

The damping term goes to zero at convergence.

D.2.2 Final Form

The final form of the Poisson equation, both for the inner points as well as the boundary points, can be easily written to be:

$$A_W\phi_W-A_P\phi_P+A_E\phi_E=Q_P \tag{D.20}$$

Where the coefficients can be shown to posses the following property of diagonal dominance:

$$A_P \geq A_W + A_E \qquad \text{And} \qquad A_E > 0 \quad A_W > 0 \quad A_P > 0 \tag{D.21}$$

These equations can be put into a matrix form

$$\bar{\bar{A}}\bar{\phi} = \bar{Q} \tag{D.22}$$

The matrix $\bar{\bar{A}}$ is tridiagonal for a one-dimensional problem; it can be solved efficiently by the Thomas' tridiagonal matrix algorithm (TDMA) [113, 114].

APPENDIX E

HOLE-CONTINUITY EQUATION

E.1 The Boundary-Value Problem

The drift-diffusion hole-continuity equation can be written by equating the divergence of the flux to generation.

$$\nabla \cdot \mathbf{J}_p = -qR \tag{E.1}$$

Where R is the net recombination, and the hole current density, $\mathbf{J}_p$, is the sum of the drift and diffusion components:

$$\mathbf{J}_p = qp\mu_p\mathbf{E} - qD_p\nabla p \tag{E.2}$$

p is the hole concentration; μ_p is the hole mobility; the diffusivity of holes is $D_p = \mu_p V_t$; and $V_t = k_B T_L/q$ is the thermal voltage.

Recombination R is taken to be the standard Shockley-Hall-Read formula (Selberherr [72] p. 105):

$$R = R_{\mathrm{SHR}} = \frac{np - n_i^2}{\tau_p(n + n_i) + \tau_n(p + n_i)} \tag{E.3}$$

where n_i is the intrinsic concentration, and the time constants, τ_n and τ_p, have the form (p. 106 [72])

$$\tau_n = \frac{\tau_{n0}}{1 + \dfrac{|\mathcal{D}|}{N_n^{\mathrm{ref}}}} \quad \text{and} \quad \tau_p = \frac{\tau_{p0}}{1 + \dfrac{|\mathcal{D}|}{N_p^{\mathrm{ref}}}} \tag{E.4}$$

Where $\mathcal{D} = N_D - N_A$ is the net doping concentration, and the constants take on the values as follows: $\tau_{n0} = 3.94 \times 10^{-4}$ sec, $\tau_{p0} = 3.94 \times 10^{-5}$ sec, $N_n^{\mathrm{ref}} = 7.1 \times 10^{15}$ cm^{-3}, $N_p^{\mathrm{ref}} = 7.1 \times 10^{15}$ cm^{-3}. And the mobility is given by

$$\mu_p = \frac{\mu_p^{(0)}}{\sqrt{1 + \dfrac{|\mathcal{D}|}{C_p^{\mathrm{ref}} + \frac{|\mathcal{D}|}{S_P}}}} \tag{E.5}$$

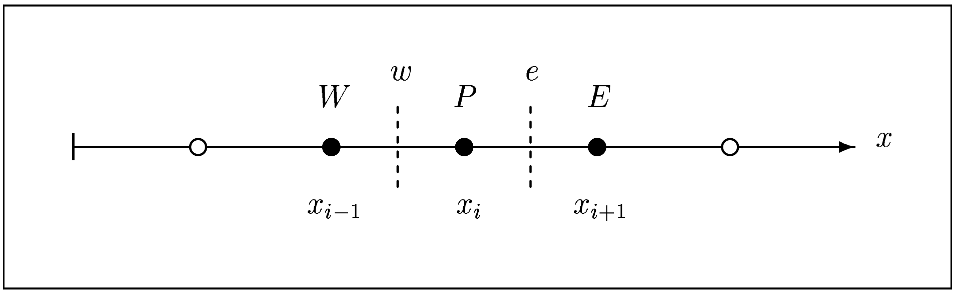


Figure E.1: Grid for the discretization of the hole-continuity equation. Filled circles indicate the computational molecule.

And the constants have values $C_p^{\text{ref}} = 4 \times 10^{16}$ cm^{-3}, and $S_p = 81$ (dimensionless) $\mu_p^{(0)} = 480$ cm^2/(V sec)

E.1.1 Boundary Conditions

To complete the description for the hole-continuity boundary-value problem we need boundary conditions. At the ohmic contact we can assume a Dirichlet boundary condition, that is, the hole concentration is specified to be equal to the equilibrium hole concentration. By charge neutrality at the ohmic boundary we have

$$p - n + \mathcal{D} = 0 \tag{E.6}$$

using the law of mass action, $np = n_i^2$, we get the boundary condition for the hole-continuity equation

$$p = \frac{-\mathcal{D} + \sqrt{\mathcal{D}^2 + 4n_i^2}}{2} \tag{E.7}$$

Eq. (E.7) is convenient to use if $\mathcal{D} < 0$ (p-type doping). For $\mathcal{D} > 0$ (n-type doping) we can use

$$p = \frac{n_i^2}{n} \qquad \text{where} \qquad n = \frac{\mathcal{D} + \sqrt{\mathcal{D}^2 + 4n_i^2}}{2} \tag{E.8}$$

E.2 Scharfetter-Gummel Discretization

We discretize the hole-continuity equation by the celebrated Scharfetter-Gummel discretization [100]. Let us consider the one-dimensional hole-continuity equation

$$\frac{dJ}{dx} + qR = 0 \tag{E.9}$$

J is the hole current density, the subscript "$_p$" has been dropped for clarity. Integrating this equation from x_w to x_e over the control volume shown in Fig. E.1 on the facing page:

$$\int_w^e \frac{dJ}{dx}\,dx + \int_w^e qR\,dx = 0 \tag{E.10}$$

The first integral is straight-forward:

$$\int_w^e \frac{dJ}{dx}\,dx = J_e - J_w \tag{E.11}$$

The second integral is approximated by assuming that that the nodal value of R at point P dominates, and it is approximately the same as the average of R around P. This approximation is second-order accurate [114].

$$\int_w^e qR\,dx = qR_P \Delta x_P + \mathcal{O}\left((\Delta x_P)^2\right) \tag{E.12}$$

where $\Delta x_P = x_e - x_w = (x_E - x_W)/2$. By substituting in Eq. (E.10) the two integrals we get the second-order accurate central-difference approximation of the hole-continuity equation

$$\frac{J_e - J_w}{\Delta x_P} + qR_P = 0 \tag{E.13}$$

We now need discretized approximations for J_e and J_w to be substituted in Eq. (E.13). With this in mind Eq. (E.2) for current density J in one dimension reads

$$\begin{aligned} J &= qp\mu_p E_x - qD_p \frac{dp}{dx} \\ &= -qp\mu_p \frac{d\phi}{dx} - q\mu_p V_t \frac{dp}{dx} \end{aligned} \tag{E.14}$$

If we assume that $J = J_e$ and $E_x = -d\phi/dx$ are constant in the interval $x \in [x_P, x_E]$, the above equation is then a in-homogeneous ordinary differential equation in p with $p(x_P) = p_P$ as the boundary condition, whose solution is

$$p(x) = \left(p_P + \frac{J_e}{q\mu_p \dfrac{d\phi}{dx}}\right) \exp\left(-\frac{\phi(x) - \phi_P}{V_t}\right) - \frac{J_e}{q\mu_p \dfrac{d\phi}{dx}} \tag{E.15}$$

Setting $x = x_E$ and $p_E = p(x_E)$ we can solve for the constant J_e

$$J_e = \frac{q\mu_p^e V_t}{\Delta x_e}\left(p_P B(\alpha^+) - p_E B(-\alpha^+)\right) \tag{E.16}$$

where $\Delta x_e = x_E - x_P$ and $\alpha^+ = (\phi_E - \phi_P)/V_t$ is the argument of the Bernoulli function B:

$$B(\Delta) \triangleq \frac{\Delta}{\exp(\Delta) - 1} \tag{E.17}$$

Similar expression can be written for J_w

$$J_w = \frac{q\mu_p^w V_t}{\Delta x_w}\big(p_W B(\alpha^-) - p_P B(-\alpha^-)\big) \tag{E.18}$$

where $\Delta x_w = x_P - x_W$ and $\alpha^- = (\phi_P - \phi_W)/V_t$.

By substituting the expression for J_e, Eq. (E.16), and J_w, Eq. (E.18), into Eq. (E.13) we get

$$\frac{qV_t}{\Delta x_P}\left(\mu_p^e \frac{p_P B(\alpha^+) - p_E B(-\alpha^+)}{\Delta x_e} - \mu_p^w \frac{p_W B(\alpha^-) - p_P B(-\alpha^-)}{\Delta x_w}\right) + qR_P = 0 \tag{E.19}$$

The recombination term R_P, given by Eq. (E.3), is split into two parts: The constant known part is moved to the right-hand side; and the unknown portion, involving p_P, is added to the diagonal term, which enhances diagonal-dominance. This splitting is known as the Seidman modification. The system of equations in matrix notation are

$$\bar{\bar{A}}\bar{p} = \bar{Q} \tag{E.20}$$

The coefficient matrix $\bar{\bar{A}}$ is tridiagonal, and can be easily and efficiently solved by the Thomas' tridiagonal matrix algorithm (TDMA) [113, 114].

APPENDIX F

LIST OF SYMBOLS

Symbol	**Meaning**
C	Pre-Maxwellian variable
$c_{\mathrm{ac}}^{\mathrm{inelastic}}$	Coefficient in expression for inelastic acoustic phonon
$c_{\mathrm{ac,\ gC}}^{\mathrm{inelastic}}$	Coefficient in expression for inelastic acoustic phonon in pre-Maxwellian variable
c_g	Coefficient in definition of $g(\varepsilon)$ $(= 4\pi\sqrt{2}(m^*)^{3/2}/\hbar^3)$
c_{ib}	Coefficient in expression for inter-band optical phonon
c_{impu}	Coefficient in expression for ionized impurity scattering
c_j	Coefficient in expression for current $(= 1/24\pi^3\sqrt{\pi})$
c_{opt}	Coefficient in expression for optical phonon
$\mathcal{D}$	Net doping $(= N_D^+ - N_A^-)$
$\mathcal{D}_{\mathrm{ac}}$	Deformation potential for acoustic phonon
$\mathcal{D}_{\mathrm{ib}}$	Deformation potential for inter-band optical phonon
$\mathcal{D}_{\mathrm{opt}}$	Deformation potential for optical phonon
$\mathbf{E}$	Electric field
E	Electric-field magnitude $(= \lvert\mathbf{E}\rvert)$
$F_l^m(H)$	Coefficient of spherical harmonic Y_l^m
$f(\mathbf{r}, \mathbf{k}, t)$	Distribution function
$f_l^m(\varepsilon)$	Coefficient of spherical harmonic Y_l^m
G_{impact}	Generation due to impact ionization
$g(\varepsilon)$	Function proportional to density of states $(= c_g\gamma'(\varepsilon)\sqrt{\gamma(\varepsilon)})$
$g^+(\varepsilon)$	$g(\varepsilon + \hbar\omega_{\mathrm{opt}})$
$g^-(\varepsilon)$	$g(\varepsilon - \hbar\omega_{\mathrm{opt}})$
H	Hamiltonian
h	Planck's constant
$\hbar$	Reduced Planck's constant $(= h/2\pi)$
$\hbar\omega_{\mathrm{ib}}$	Energy of inter-band optical phonon
$\hbar\omega_{\mathrm{opt}}$	Energy of optical phonon

Symbol	**Meaning**
I_{DS}	Drain-to-source current
$\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$	Unit vectors in Cartesian directions x, y, z
i, j, k	Index for grid (x_i, y_j, H_k)
$\mathbf{J}$	Current density vector
J	Current density magnitude ($= \lvert\mathbf{J}\rvert$)
J_x	x-component of current density $\mathbf{J}$
$\mathbf{k}$	Wave-vector
k	Wave-vector magnitude ($= \lvert\mathbf{k}\rvert$)
k, θ, ϕ	Spherical component of wave-vector $\mathbf{k}$
k_B	Boltzmann constant
k_x, k_y, k_z	Cartesian components of wave-vector $\mathbf{k}$
L	Effective channel length of MOSFET
L_{gate}	Gate length of MOSFET
$L_{\text{metallurgical}}$	Metallurgical channel length of MOSFET
m^*	Effective mass
N_A	Number of acceptors per unit volume
N_A^-	Number of ionized acceptors per unit volume
N_D	Number of donors per unit volume
N_D^+	Number of ionized donors per unit volume
N_I	Number of impurities per unit volume
N_{opt}	Number of optical phonons
N_q	Number of phonons
n	Electron concentration
n_0	Equilibrium electron concentration
n_i	Intrinsic concentration
P_{impact}	Impact ionization scattering pre-factor
$\mathbf{p}$	Quantum-mechanical momentum of electron
p	Hole concentration
$\mathbf{q}$	Wave-vector of phonon
q	Magnitude of electronic charge
q_D	Reciprocal of Debye length
$\mathbf{r}$	Real-space vector
r_e	Expansion factor for MOSFET grid generation
$S(\mathbf{k}, \mathbf{k}')$	Transition rate from $\mathbf{k}$ to $\mathbf{k}'$
T_L	Lattice temperature
t	Time
t_{ox}	Oxide thickness in MOSFET
u	Slotboom variable
$\mathbf{u}_g$	Group velocity
u_g	Group-velocity magnitude ($= \lvert\mathbf{u}_g\rvert$)
um	Micron (μm), or micro-meter ($1\mu\text{m} = 10^{-6}\text{m}$)

Symbol	**Meaning**
v	Slotboom variable
V_{BE}	Base-to-emitter voltage
V_{CB}	Collector-to-base voltage
V_{CE}	Collector-to-emitter voltage
V_{DS}	Drain-to-source voltage
V_{GS}	Gate-to-source voltage
V_T	Threshold voltage of MOSFET
V_t	Thermal voltage ($= k_B T_L / q$)
v_{sound}	Velocity of sound
W	Width of MOSFET
x, y, z	Cartesian space directions
x	Cartesian space direction: position for n^+nn^+; depth for BJT; lateral direction for MOSFET
x_{overlap}	Oxide overlap with source/drain in MOSFET
y	Cartesian space direction: depth direction for MOSFET
y_{junction}	Source/Drain junction-depth in MOSFET
$Y_l^m(\theta, \phi)$	Spherical-harmonic function
Z	Number of charge units of the impurity
$Z^{(\nu)}$	Band multiplicity (number of equivalent symmetrical bands) of band (ν)
α	Non-parabolicity factor
α_{impact}	Impact ionization coefficient
$\gamma(\varepsilon)$	Dispersion relation
ΔH	Grid spacing Hamiltonian
Δx	Grid spacing in x
Δy	Grid spacing in y
ε	Energy of electron
ϵ_0	Perimittivity of free space
ϵ_{rel}	Relative permittivity of Silicon
ϵ_{si}	Perimittivity of Silicon
θ	Polar angle
κ	"Conductivity" in Boltzmann equation ($= \tau u_g^2 g$)
μm	Micron, or micro-meter ($1\mu\text{m} = 10^{-6}\text{m}$)
ν	Band index in multi-band band-structure
ξ_{impurity}	Impurity scattering enhancement factor
ρ	Density of silicon
σ	Standard deviation of Gaussian doping in MOSFET
τ	Reciprocal of scattering rate
τ_{ac}	Reciprocal of acoustic phonon scattering rate
$\tau_{\text{ib}}^{(\nu \leftrightarrow \nu')}$	Reciprocal of inter-band optical phonon scattering rate

Symbol	**Meaning**
τ_{impact}	Reciprocal of impact ionization scattering rate
$\tau_{l=1,\,\text{impurity}}$	Reciprocal of ionized impurity scattering rate
τ_{opt}	Reciprocal of optical phonon scattering rate
ϕ	Azimuthal angle
$\phi(\mathbf{r})$	Electric potential
Ω	Volume of semiconductor crystal sample ($= L_x L_y L_z$)
ω	Relaxation factor for SOR (Successive Over-Relaxation)
ω_{ib}	Frequency of inter-band optical phonon
ω_{opt}	Frequency of optical phonon
ω_q	Frequency of phonon
∇	Nabla (or del) operator
∇_k	Nabla (or del) operator in wave-vector space
∇_r	Nabla (or del) operator in real space

BIBLIOGRAPHY

Introduction

[1] T. Forester, *Silicon Samurai: How Japan Conquered the world's I.T. Industry*, Blackwell Publishers, Cambridge, USA, pp. 5–6 & 43, 1993.

[2] The National Technology Roadmap for Semiconductors, `http://notes.sematech.org/97melec.htm` in the Sematech website at `http://www.sematech.org/public/roadmap`, 1997.

[3] R. J. G. Goossens, and R. W. Dutton, "Device CAD in the '90s: At the Crossroads," *IEEE Circuits & Devices Magazine*, Vol. 8, No. 4, pp 18–26, 1992.

[4] M. E. Law, "The virtual IC Factory ... Can it be Achieved ?" *IEEE Circuits & Devices Magazine*, Vol. 11, No. 2, pp 25–31, 1995.

[5] P. A. Blakey and T. E. Zirkle, "An Industrial Perspective on Semiconductor Technology Modeling," *The IMA Volumes in Mathematics and its Applications, Semiconductors, Part II*, Edited by W. M. Coughran, J. Cole, P. LLoyd, and J. White, Vol. 59, pp. 75–87, Springer-Verlag, 1994.

[6] P. Packan, "Simulating Deep Sub-Micron Technologies: An Industrial Perspective," *Proc. SISDEP*, Vol. 6, pp. 34–41, September 1995.

[7] H. Kosina, E. Langer, and S. Selberherr, "Device Modelling for the 1990s," *Microelectronics Journal*, Vol. 26, pp. 217–233, 1995.

[8] J. W. Specks and W. L. Engl, "Computer-Aided Design and Scaling of Deep Submicron CMOS," *IEEE Trans. Computer-Aided Design*, Vol. 12, No. 9, pp. 1357–1367, 1993.

[9] P. Rohr, F. A. Lindholm, and K. R. Allen, "Questionability of Drift-Diffusion Transport in the Analysis of Small Semiconductor Devices," *Solid-State Electron.*, Vol. 17, pp. 729–734, 1974.

[10] G. Baccarani, F. Odeh, A. Gnudi, and D. Ventura, “A Critical Review of the Fundamental Semiconductor Equations,” *The IMA Volumes in Mathematics and its Applications, Semiconductors, Part II*, Edited by W. M. Coughran, J. Cole, P. LLoyd, and J. White, Vol. 59, pp. 19–32, Springer-Verlag, 1994.

[11] Refer to the web site at `http://www-groups.cs.st-and.ac.uk/~history/Mathematicians/Boltzmann.html`

Basic Device Simulation

[12] A. Abramo *et al.*, “A Comparison of Numerical Solutions of the Boltzmann Transport Equation for High-Energy Electron Transport Silicon,” *IEEE Trans. Electron Devices*, Vol. 41, No. 9, pp. 1646–1654, 1994.

[13] M. A. Alam, M. A. Stettler, and M. S. Lundstrom, “Formulation of the Boltzmann Equation in Terms of Scattering Matrices,” *Solid-State Electron.*, Vol. 36, No. 2, pp. 263–271, 1993.

[14] F. J. Mustieles and F. Delaurens, “Numerical Simulation of Non-Homogeneous Submicron Semiconductor Devices by a Deterministic Particle Method,” *Solid-State Electron.*, Vol. 36, No. 6, pp. 857–868, 1993.

[15] G. Zandler, A. D. Carlo, K. Kometer, P. Lugli, P. Vogl, and E. Gornik, “A Comparison of Monte Carlo and Cellular Automata Approaches for Semiconductor Device Simulation,” *IEEE Electron Device Lett.*, Vol. 14, No. 2, pp. 77–79, 1993.

[16] B. H. Floyd and Y. L. Le Coz, “Iterative Spectral Solution of the Poisson-Boltzmann Equation in Semiconductor Devices,” *J. Appl. Phys.*, Vol. 76, No. 12, pp. 7889–7898, 1994.

Spherical Harmonics

[17] K. Rahmat, J. White, and D. A. Antoniadis, “Simulation of Semiconductor Devices Using a Galerkin/Spherical Harmonic Expansion Approach to Solving the Coupled Poisson-Boltzmann System,” *IEEE Trans. Computer-Aided Design*, Vol. 15, No. 10, pp. 1181–1196, 1996.

[18] K. Rahmat, J. White, and D. A. Antoniadis, “Solution of the Boltzmann Transport Equation in Two Real-Space Dimensions using a Spherical Harmonic Expansion in Momentum Space,” *IEDM-1994 Tech. Dig.*, pp. 359–362, 1994.

[19] K. Rahmat, J. White, and D. A. Antoniadis, "A Galerkin Method for the Arbitrary Order Expansion in Momentum Space of the Boltzmann Equation using Spherical Harmonics," *Proc. of the NUPAD V Conf.*, Honolulu, HI, pp. 133–136, June 1994.

[20] N. Goldsman, L. Henrickson, and J. Frey, "A Physics Based Analytical/Numerical Solution to the Boltzmann Transport Equation for Use in Device Simulation," *Solid-State Electron.*, Vol. 34, No. 4, pp. 389–396, 1991.

[21] D. Ventura, A. Gnudi, and G. Baccarani, "An Efficient Method for Evaluating the Energy Distribution of Electrons in Semiconductors Based on Spherical Harmonics Expansion," *IEICE Trans. Electron.*,Vol. E75-C, No. 2, pp. 194–199, 1992.

[22] H. Lin, and N. Goldsman, "An Efficient Solution of the Boltzmann Transport Equation which Includes the Pauli Exclusion Principle," *Solid-State Electron.*, Vol. 34, No. 10, pp. 1035–1047, 1991.

[23] Y.-J. Wu and N. Goldsman, "An Efficient Solution of the Multi-Band Boltzmann Transport Equation in Silicon," *COMPEL*, Vol. 12, No. 4, pp. 475–485, 1993.

[24] N. Goldsman, Y.-J. Wu, and J. Frey, "Efficient Calculation of Ionization Coefficients in Silicon from the Energy Distribution Function," *J. Appl. Phys.*, Vol. 68, No. 3, pp. 1075–1081, 1990.

[25] Y.-J. Wu and N. Goldsman, "Deterministic Modeling of Impact Ionization with a Random-k Approximation and the Multiband Boltzmann Equation," *J. Appl. Phys.*, Vol. 78, No. 8, pp. 5174–5176, 1995.

[26] S. L. Wang, N. Goldsman, and K. Hennacy, "Calculation of Impact Ionization Coefficients with a Third-Order Legendre Polynomial Expansion of the Distribution Function," *J. Appl. Phys.*, Vol. 68, No. 3, pp. 1815–1822, 1992.

[27] M. C. Vecchi and M. Rudan, "Modeling Impact-Ionization in the Framework of the Spherical-Harmonic Expansion of the Boltzmann Transport Equation with Full-Band Structure Effects," *Proc. SISDEP*, Vol. 6, *Edited by H. Ryssel and P. Pichler*, pp. 416–419, September 1995.

[28] M. C. Vecchi, D. Ventura, A. Gnudi, and G. Baccarani, "Incorporating Full Band-Structure Effects in the Spherical Harmonics Expansion of the Boltzmann Transport Equation," *Proc. of the NUPAD V Conf.*, Honolulu, HI, pp. 55–58, June 1994.

[29] M. C. Vecchi and M. Rudan, "Modeling Electron and Hole Transport with Full-Band Structure Effects by Means of the Spherical-Harmonic Expansion of the BTE," *IEEE Trans. Electron Devices*, Vol. 45, No. 1, pp. 230–238, 1998.

[30] D. Ventura, A. Gnudi, and G. Baccarani, "Inclusion of Electron-Electron Scattering in the Spherical Harmonic Expansion Treatment of the Boltzmann Transport Equation," *Proc. SISDEP*, Vol. 5, *Edited by S. Selberherr, H. Stippel, and E. Strasser*, pp. 161–164, 1993.

[31] A. Gnudi, D. Ventura, G. Baccarani, and F. Odeh, "Macroscopic and Microscopic Approach for the Simulation of Short Devices," *The IMA Volumes in Mathematics and its Applications, Semiconductors, Part II*, Edited by W. M. Coughran, J. Cole, P. LLoyd, and J. White, Vol. 59, pp. 135–157, Springer-Verlag, 1994.

[32] H. Lin, N. Goldsman, and I. D. Mayergoyz, "An Efficient Deterministic Solution of the Space-Dependent Boltzmann Transport Equation for Silicon," *Solid-State Electron.*, Vol. 35, No. 1, pp. 33–42, 1992.

[33] H. Lin, N. Goldsman, and I. D. Mayergoyz, "Device Modeling by Deterministic Self-Consistent Solution of Poisson and Boltzmann Transport Equations," *Solid-State Electron.*, Vol. 35, No. 6, pp. 769–778, 1992.

[34] H. Lin, N. Goldsman, and I. D. Mayergoyz, "Deterministic BJT Modeling by Self-Consistent Solution to the Boltzmann, Poisson and Hole-Continuity Equations," *Proc. International Workshop on Computational Electronics, Univ. of Leeds, England*, pp. 55–59, 1993.

[35] D. Ventura, A. Gnudi, G. Baccarani, and F. Odeh, "Multidimensional Spherical Harmonic Expansion of Boltzmann Equation for Transport in Semiconductors," *Appl. Math. Lett.*, Vol. 5, No. 3, pp. 85–90, 1992.

[36] A. Gnudi, D. Ventura, and G. Baccarani, "One-Dimensional Simulation of a Bipolar Transistor by means of Spherical Harmonics Expansion of the Boltzmann Transport Equation," *Proc. SISDEP*, Vol. 4, *Edited by W. Fichtner, D. Aemmer, '91 Conf. (Zurich)*, pp. 205–213, September 1991.

[37] A. Gnudi, D. Ventura, and G. Baccarani, "Modeling Impact Ionization in a BJT by Means of Spherical Harmonics Expansion of the Boltzmann Transport Equation," *IEEE Trans. Computer-Aided Design*, Vol. 12, No. 11, pp. 1706–1713, 1993.

[38] D. Schroeder, D. Ventura, A. Gnudi, and G. Baccarani, “Boundary Conditions for Spherical Harmonics Expansion of Boltzmann Equation,” *Electronics Letters*, Vol. 28, No. 11, pp. 995–996, 1992.

[39] A. Gnudi, D. Ventura, and G. Baccarani, “Two-Dimensional MOSFET Simulation by Means of a Multidimensional Spherical Harmonic Expansion of the Boltzmann Transport Equation,” *Solid-State Electron.*, Vol. 36, No. 4, pp. 575–581, 1993.

[40] K. A. Hennacy and N. Goldsman, “A Generalized Legendre Polynomial/Sparse Matrix approach for determining The Distribution Function in Non-Polar Semiconductors,” *Solid-State Electron.*, Vol. 36, No. 6, pp. 869–877, 1993.

[41] K. A. Hennacy, Y. -J. Wu, N. Goldsman, and I. D. Mayergoyz, “ Deterministic MOSFET Simulation Using a Generalized Spherical Harmonic Expansion of the Boltzmann Equation,” *Solid-State Electron.*, Vol. 38, No. 8, pp. 1485–1495, 1995.

[42] W.-C. Liang, Y.-J. Wu, K. Hennacy, S. Singh, N. Goldsman, and I. Mayergoyz, “2-D MOSFET Simulation by Self-Consistent Solution of the Boltzmann and Poisson Equations Using a Generalized Spherical Harmonic Expansion,” *Proc. SISDEP*, Vol. 6, *Edited by Ryssel and P. Pichler*, pp. 122–125, 1995.

[43] W. Liang, N. Goldsman, I. D. Mayergoyz, and P. J. Oldiges, “2-D MOSFET Modeling Including Surface Effects and Impact Ionization by Self-Consistent Solution of the Boltzmann, Poisson, and Hole-Continuity Equations,” *IEEE Trans. Electron Devices*, Vol. 44, No. 2, pp. 257–267, 1997.

[44] C.-H. Chang, C.-K. Lin, W. Liang, N. Goldsman, I. D. Mayergoyz, P. Oldiges, and J. Melngailis, “The Spherical Harmonic Method: Corroboration with Monte Carlo and Experiment,” Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD), 1997, Cambridge, MA, USA, pp. 225–228, 1997.

[45] A. Pierantoni, M. C. Vecchi, and A. Gnudi, “Sub-Domain Solution of the Boltzmann Equation in MOS Devices by Means of Spherical Harmonic Expansion,” Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD), 1997, Cambridge, MA, USA, pp. 229–232, 1997.

[46] M. C. Vecchi, J. Mohring, and M. Rudan, “An Efficient Solution Scheme for the Spherical-Harmonic Expansion of the Boltzmann Transport Equa-

tion," *IEEE Trans. Computer-Aided Design*, Vol. 16, No. 4, pp. 353–361, 1997.

PhD Theses

[47] N. Goldsman, *Modeling Electron Transport and Degradation Mechanisms in Semiconductor Submicron Devices*, PhD Thesis, Cornell University, Ithaca, NY, 1989.

[48] H. C. Lin, *Efficient Self-Consistent Semiconductor Device Modeling by Deterministic Solution to the Boltzmann Transport Equation*, PhD Thesis, University of Maryland, College Park, MD, 1992.

[49] K. A. Hennacy, *Spherical Harmonics and Effective Field Formulations of Boltzmann's Transport Equation: Case Studies in Silicon*, PhD Thesis, University of Maryland, College Park, MD, 1993.

[50] Y.-J. Wu, *Multi-Band and Two Dimensional Submicron Semiconductor Device Modeling by Direct Solution to the Boltzmann Transport Equation*, PhD Thesis, University of Maryland, College Park, MD, 1994.

[51] W. Liang, *Two-Dimensional Submicron Semiconductor Device TCAD by Hydrodynamic and Numerical Boltzmann Simulation*, PhD Thesis, University of Maryland, College Park, MD, 1996.

[52] X. Wang, *SLAPSHOT: An Engineering Oriented Monte Carlo Tool for Hot Carrier Studies in Deep Submicron MOSFETs*, PhD Thesis, University of Texas, Austin, TX, 1994.

[53] A. W. Duncan, *Full-Band Monte Carlo Simulation of Hot Electrons in Scaled Silicon Devices*, PhD Thesis, University of Illinois, Urbana-Champaign, IL, 1996.

[54] M. Liang, *Composite Thermodynamic System for Advanced Semiconductor Device Simulation*, PhD Thesis, University of Florida, Gainesville, FL, 1994.

[55] D. C. Kerr, *Three-Dimensional Drift-Diffusion and Hydrodynamic Simulation of Semiconductor Devices*, PhD Thesis, University of Maryland, College Park, MD, 1995.

[56] C.-F. Yeap, *UT-MiniMOS: A Hierarchical Transport Model Based Simulator for Deep Submicron Silicon Devices*, PhD Thesis, University of Texas, Austin, TX, 1997.

Band Structure & Scattering

[57] R. Brunetti and C. Jacoboni, F. Venturi, E. Sangiorgi, and B. Riccò, “A Many-Band Silicon Model for Hot-Electron Transport at High Energies,” *Solid-State Electron.*, Vol. 32, No. 12, pp. 1663–1667, 1989.

[58] C. Fiegna and E. Sangiorgi, “Modeling of High-Energy Electrons in MOS Devices at the Microscopic Level,” *IEEE Trans. Electron Devices*, Vol. 40, No. 3, pp. 619–627, 1993.

[59] J. Kolodziejczak, “On the Scattering Process in Semiconductors,” *Phys. Stat. Sol.*, Vol. 19, pp. 231–237, 1967.

[60] R. Thoma, H. J. Peifer, W. L. Engl, W. Quade, R. Brunetti, and C. Jacoboni, “An Improved Impact-Ionization Model for High-Energy Electron Transport in Si with Monte Carlo Simulation,” *J. Appl. Phys.*, Vol. 69, No. 4, pp. 2300–2311, 1991.

[61] F. Venturi, E. Sangiorgi, R. Brunetti, C. Jacoboni, and B. Riccò, “Monte Carlo Simulation of Electron Heating in Scaled Deep Submicron Mosfets,” *IEDM-1989 Tech. Dig.*, pp. 485–488, 1989.

Monte Carlo

[62] K. Tomizawa, *Numerical Simulation of Submicron Semiconductor Devices*, Artech House, Boston, MA, 1993.

[63] C. Moglestue, *Monte Carlo Simulation of Semiconductor Devices*, Chapman & Hall, London, 1993.

[64] C. Jacoboni, and P Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*, Springer-Verlag Wien, New York, 1989.

[65] K. Hess, *Monte Carlo Simulation: Full Band and Beyond*, Kluwer Academic Publishers, Norwell, MA, 1991.

[66] C. Jacoboni and L. Reggiani, “The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials,” *Reviews of Modern Physics*, Vol. 55, No. 3, pp. 645–705, 1983.

[67] P. Lugli, “The Monte Carlo Method for Semiconductor Device and Process Modeling,” *IEEE Trans. Computer-Aided Design*, Vol. 9, No. 11, pp. 1164–1176, 1990.

[68] D. L. Woolard, H. Tian, M. A. Littlejohn, and K. W. Kim, “The Implementation of Physical Boundary Conditions on the Monte Carlo Simulation of Electron Devices,” *IEEE Trans. Computer-Aided Design*, Vol. 13, No. 10, pp. 1241–1246, 1994.

[69] P. Lugli, “Monte Carlo Models and Simulations,” in *Compound Semiconductor Device Modelling*, Edited by C. M. Snowden and R. E. Miles, Springer-Verlag, London, 1993.

[70] F. Venturi, R. K. Smith, E. C. Sangiorgi, M. R. Pinto, and B. Riccó, “A General Purpose Device Simulator Coupling Poisson and Monte Carlo Transport with Application to Deep Submicron MOSFET’s,” *IEEE Trans. Computer-Aided Design*, Vol. 8, No. 4, pp. 360–369, 1989.

Books

[71] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, New York, 1985.

[72] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag Wien, New York, 1984.

[73] S. Datta, *Quantum Phenomena*, Modular Series on Solid State Devices, V. 8, Addison-Wesley Publishing Company Inc., Reading, MA, 1989.

[74] M. Lundstrom, *Fundamentals of Carrier Transport*, Modular Series on Solid State Devices, V. X, Addison-Wesley Publishing Company Inc., Reading, MA, 1990.

[75] B. K. Ridley, *Quantum Processes in Semiconductors*, 3rd ed., Oxford University Press, Oxford, United Kingdom, 1993.

[76] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., John Wiley & Sons, New York, 1981.

[77] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser, *Semiconductor Equations*, Springer-Verlag Wien, New York, 1990.

[78] S. M. Sze, Editor, *High-Speed Semiconductor Devices*, John Wiley & Sons Inc., New York, 1990.

[79] D. Schroeder, *Modelling of Interface Carrier Transport for Device Simulation*, Springer-Verlag Wien, New York, 1994.

[80] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, John Wiley & Sons Inc., New York, second edition, 1986.

[81] R. W. Dutton and Z. Yu, *Technology CAD—Computer Simulation of IC Processes and Devices*, Kluwer Academic Publishers, Norwell, MA, 1993.

[82] K. M. Cham, S.-Y. Oh, D. Chin, J. L. Moll, K. Lee, and P. V. Voorde, *Computer-Aided Design and VLSI Device Development*, Second edition, Kluwer Academic Publishers, Boston, MA, 1988.

[83] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, Oxford University Press, New York, 1987.

[84] N. Arora, *MOSFET Models for VLSI Circuit Simulation Theory and Practice*, Springer-Verlag Wien, New York, 1993.

[85] Y. Leblebici and S. M. Kang, *Hot-Carrier Reliability of MOS VLSI Circuits*, Kluwer Academic Publishers, Norwell, MA, 1993.

Bipolar Junction Transistor

[86] A. Chryssafis and W. Love, "A Computer-aided Analysis of One-dimensional Thermal Transients in n-p-n Power Transistors," *Solid-State Electron.*, Vol. 22, pp. 249–256, 1979.

[87] E. F. Crabbé, J. M. C. Stork, G. Baccarani, M. V. Fischetti, and S. E. Laux, "The Impact of Non-Equilibrium Transport on Breakdown and Transit time in Bipolar Transistors," *IEDM-90 Tech. Dig.*, pp. 463–466, 1990.

[88] G. Baccarani, C. Jacoboni, and A. M. Mazzone, "Current Transport in Narrow-Base Transistors," *Solid-State Electron.*, Vol 20, pp. 5–10, 1977.

[89] Y.-J. Park, D. H. Navon, and T. W. Tang, "Monte Carlo Simulation of Bipolar Transistors," *IEEE Trans. Electron Devices*, Vol. ED-31, No. 12, pp. 1724–1730, 1984.

[90] W. Lee, S. E. Laux, M. V. Fischetti, and D. D. Tang, "Monte Carlo Simulation of Non-Equilibrium Transport in Ultra-Thin Base Si Bipolar Transistors," *IEDM-89 Tech. Dig.*, pp. 473–476, 1989.

[91] M. A. Stettler and M. S. Lundstrom, "A Microscopic Study of Transport in Thin Base Silicon Bipolar Transistors," *IEEE Trans. Electron Devices*, Vol. 41, No. 6, pp. 1027–1033, 1994.

[92] S. E. Laux, M. V. Fischetti, and D. J. Frank, "Monte Carlo Analysis of Semiconductor Devices: The DAMOCLES Program," *IBM J. Res. Develop.*, Vol. 34, No. 4, pp. 466–493, 1990.

[93] W. Lee, S. E. Laux, M. V. Fischetti, G. Baccarani, A Gnudi, J. M. C. Stork, J. A. Mandelman, E. F. Crabbé, M. R. Wordeman, and F. Odeh, "Numerical Modeling of Advanced Semiconductor Devices," *IBM J. Res. Develop.*, Vol. 36, No. 2, pp. 208–232, 1992.

[94] A. D. Carlo and P. Lugli, "Dead-Space Effects Under Near-Breakdown Conditions in AlGaAs/GaAs HBT's," *IEEE Electron Device Lett.*, Vol. 14, No. 3, pp. 103–106, 1993.

Device Simulation

[95] I. D. Mayergoyz, "Solution of Nonlinear Poisson Equation of Semiconductor Device Theory," *J. Appl. Phys.*, Vol. 59, No. 1, pp. 195–199, 1986.

[96] G. J. L. Ouwerling, "Further Improved algorithm for the Solution of the Nonlinear Poisson Equation in Semiconductor Devices," *J. Appl. Phys.*, Vol. 66, No. 12, pp. 6144–6149, 1989.

[97] H. Kobeissi, F. M. Ghannouchi, and A. Khebir, "Finite Element Solution of the Nonlinear Poisson Equation for Semiconductor Devices using the Fixed-Point Iteration Method," *J. Appl. Phys.*, Vol. 74, No. 10, pp. 6186–6190, 1993.

[98] C. E. Korman and I. D. Mayergoyz, "A Globally Convergent Algorithm for the Solution of the Steady-State Semiconductor Device Equations," *J. Appl. Phys.*, Vol. 68, No. 3, pp. 1324–1334, 1990.

[99] H. K. Gummel, "A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculation," *IEEE Trans. Electron Devices*, Vol. 11, No. 10, pp. 455–465, 1964.

[100] D. L. Scharfetter and H. K. Gummel, "Large-Signal Analysis of a Silicon Read Diode Oscillator," *IEEE Trans. Electron Devices*, Vol. 16, No. 1, pp. 64–77, 1969.

[101] R. E. Bank, W. M. Coughran, M. A. Driscoll, R. K. Smith, and W. Fichtner, "Iterative Methods in Semiconductor Device Simulation," *Computer Physics Communications*, Vol. 53, pp. 201–212, North-Holland, Amsterdam, 1989.

[102] C. S. Rafferty, M. R. Pinto, and R. W. Dutton, "Iterative Methods in Semiconductor Device Simulation," *IEEE Trans. Electron Devices*, Vol. ED-32, No. 10, pp. 2018–2027, 1985.

[103] A. F. Franz, G. A. Franz, S. Selberherr, C. Ringhofer, and P. Markowich, "Finite Boxes—A Generalization of the Finite-Difference Method Suitable for Semiconductor Device Simulation," *IEEE Trans. Electron Devices*, Vol. ED-30, No. 9, pp. 1070–1082, 1983.

[104] R. E. Bank, D. J. Rose, and W. Fichtner, "Numerical Methods for Semiconductor Device Simulation," *IEEE Trans. Electron Devices*, Vol. Ed-30, No. 9, pp. 1031–1041, 1983.

[105] W. Fichtner, D. J. Rose, and R. E. Bank, "Semiconductor Device Simulation," *IEEE Trans. Electron Devices*, Vol. ED-30, No. 9, pp. 1018–1030, 1983.

[106] P. M. de Zeeuw, CWI (Centre for Mathematics and Computer Science) Netherlands, *Private Communication*, 1998.

[107] D. Sinitsky, F. Assaderaghi, M. Orshansky, J. Bokor, and C. Hu, "Velocity Overshoot of Electrons and Holes in Si Inversion Layers," *Solid-State Electron.*, Vol. 41, No. 8, pp. 1119–1125, 1997.

[108] K. Hess and G. J. Iafrate, "Theory and Applications of Near Ballistic Transport in Semiconductors," *Proc. of IEEE*, Vol. 76, No. 5, pp. 519–532, 1988.

[109] S. Ramaswamy and T. W. Tang, "Comparison of Semiconductor Transport Models Using a Monte Carlo Consistency Test," *IEEE Trans. Electron Devices*, Vol. 41, No. 1, pp. 76–83, 1994.

[110] T. J. Bordelon, V. M. Agostinelli, Jun., X.-L. Wang, C. M. Maziar, and A. F. Tasch, "Relaxation Time Approximation and Mixing of Hot and Cold Electron Populations," *Electronics Letters*, Vol. 28, No. 12, pp. 1173–1174, 1992.

[111] B. Meinerzhagen and W. L. Engl, "The Influence of the Thermal Equilibrium Approximation on the Accuracy of Classical Two-Dimensional Numerical Modeling of Silicon Submicrometer MOS Transistors," *IEEE Trans. Electron Devices*, Vol. 35, No. 5, pp. 689–697, 1988.

[112] S. Jog and V. P. Sundarsingh, "New 2D Diffusion Simulator Using Spatial Variable Transformation," *Microelectronics Journal*, Vol. 27, pp. 571–575, 1996.

Computational Fluid Dynamics

[113] S. V. Patankar, *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation, Washington D.C., 1980.

[114] J. H. Ferziger and M. Perić, *Computational Methods for Fluid Dynamics*, Springer-Verlag, Berlin Heidelberg, 1996.

[115] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zhang, *Spectral Methods in Fluid Mechanics*, Springer-Verlag, Berlin, 1987.

[116] S. V. Patankar, "A Numerical Method for Conduction in Composite Materials, Flow in Irregular Geometries and Conjugate Heat Transfer," *Sixth International Heat Transfer Conference*, Toronto, pp. 297–302, 1978.

INDEX

A

Acoustic Phonon
  Bose-Einstein relation, 29
  deformation potential, 29
  discretization, *see* Discretization
  elastic, 8, 29, 82, 91, 92
  elastic versus inelastic, 29
  elastic, justification of, 92
  frequency of, 29
  generation due to, 174, 175
  inelastic, 8, 27, 29, **29**, 79, 82, 91, 93
  inelastic, energy exchange in, 93
  inelastic, new formulation, 54
  Kolodziejczak's derivation, 29, 82
  number of, 29
  numerical noise and, 8, 29, 82, 91, **91**, 105f, 106f
  pre-Maxwellian variable, *see* Pre-Maxwellian Variable
  scattering rate, 29, 50f
  self-adjoint form, 8, 9, 54
  transition rate, 29
  transport model in, 49
  Vecchi's scheme and, 59
Active Region
  BJT, 109, 117f
  MOSFET, 36, 140, 141
  semiconductor devices, 35
Antoniadis, Prof. Dimitri, 6

B

Baccarani, Giorgio, 133
Ballistic Transport, **114**
  BJT, **114**, 127f, 128f
  distribution function in, 114, 142, 165f
  Hamiltonian conservation in, 115
  MOSFET, 142, 165f
  signature of, 114
Band-Structure, 14
  Boltzmann equation in multi-band, 46
  Brunetti's, 8, 43, **44**, 45f, 48, 133
  Brunetti's and actual DOS, 44
  concatenation, 46
  dispersion relation, *see* Dispersion Relation
  ε-k, 44, 45f
  full-band in Spherical Harmonic, 8
  multi-band, **44**
  non-parabolicity factor, 44
  spherical symmetry, 6
  transport model, *see* Transport Model
Bernoulli Function, 192
BiCMOS, 108
BJT (Bipolar Junction Transistor)
  1-D structure & doping, 110, 117f
  2-D effects, 108, 109
  2-D structure, 117f
  active region, 109, 117f
  advantages of, **108**

  - as 1-D device, 107, 108, 109
  - ballistic transport in, **114**, 127f, 128f
  - band-structure in simulation, 108
  - base boundary condition, **109**
  - base-transit time, 115
  - BiCMOS, 108
  - bipolar nature of, 107
  - breakdown voltage, 111
  - convergence problems, 107
  - current crowding, 109
  - cut-off frequency, 115
  - DD versus BTE, 114, 126f
  - dead-space effect, 112
  - disadvantages of, 108
  - distribution function, 110, 119f, 120f
  - doping, 110, 117f
  - drift-diffusion, 107, 109, 114
  - electric field, 110, 118f
  - electric potential, 110, 118f
  - electron concentration, 111, 121f
  - electron current, 112, 122f
  - electron energy, 112, 123f
  - electron-hole plasma in, 107
  - energy-transport, 107
  - graded base doping, 115, 129f
  - grid, 109, 110
  - HBT, 108
  - holes, 112, 122f
  - hot-electron effect in, 111
  - hydrodynamic, 107
  - impact ionization, 111, 112, 113, 124f
  - impact ionization coefficient, 112, 124f
  - low-level injection, 110
  - Monte Carlo simulation of, 108
  - motivation for choice, 107
  - non-self-consistent, 108
  - n-type collector, 110
  - poly-silicon emitter, 108
  - p-type collector, 110
  - retarding potential in, 108
  - scattering-matrix approach, 108
  - self-consistent simulation, 108
  - simulation time, 110
  - simulator, 85, 110
  - spherical-harmonic simulation, 6, 9, 108
  - thermal equilibrium, **115**
  - thermal equilibrium distribution, 115, 130f
  - thermal equilibrium electric field, 116, 131f
  - thermal equilibrium energy, 115, 131f
  - velocity, 111, 121f
  - velocity overshoot, 111, 115
  - why simulate, 107
- Bologna, University of, 6
- Boltzmann Transport Equation, **12**
  - analogy to heat transfer, 37, 63, 179
  - as conservation equation, 14, 61, 173, 175
  - as first-order integro-differential equation, 19
  - as integro-differential equation, 11, 14, 19
  - band concatenation, 46
  - Boltzmann-Poisson system, 9, 177, 178
  - cellular-automata approach, 11
  - comparison with drift-diffusion, 89, 114
  - coupling in Hamiltonian, 92
  - current-conservation property, **173**
  - degenerate elliptic, 39
  - derivation of, 12
  - deterministic particle method,

- 11
- diagonal dominance, *see* Diagonal Dominance
- difficulty in numerical solution of, 11, 14
- dimensionality, *see* Dimensionality
- discretization, *see* Discretization
- elastic phonon, 92
- energy boundary, 39
- in equivalent valleys, 46
- in multi-band band-structure, 46
- iterative-spectral, 11
- linearity of, 8, 54, 55, 72
- Monte Carlo, *see* Monte Carlo
- pre-Maxwellian variable, *see* Pre-Maxwellian Variable
- reducing the number of, 46
- scattering-matrix approach, 11
- semi-classical transport in, 14
- set of independent equations, 92
- spherical harmonic, *see* Spherical Harmonic
- thermal equilibrium, *see* Thermal Equilibrium
- two-dimensional, **135**

Boltzmann, Ludwig, 13f

Boltzmann-Poisson system
- analytical solution of, 9, 89, 177
- convergence of, 90, 186
- uniqueness of solution, 178

Bose-Einstein Relation
- in acoustic phonon, 29
- in optical phonon, 27

Boundary, *see* Interface

Boundary Condition, **35**
- base of BJT, **109**
- charge neutrality at ohmic contact, 36, 190
- Dirichlet, 35, 36, 89, 109, 111, 112, 135, 184, 190
- Dirichlet and diagonal dominance, 36
- Dirichlet, advantage of, 36
- discretization, 69
- energy boundary, 35, **37**, 40f, 41f, 70, 72f
- energy boundary & original Boltzmann equation, 39
- grid for, 70f, 72f
- in boundary-value problem, 19
- insulator, 35, **36**, 70, 135
- mixed, 35, 39
- MOSFET, **135**
- Neumann, 70f, 184
- ohmic, 35, **35**, 70, 89, 109, 111, 184, 190
- Schottky, 35n
- substitute, 37, 39
- zero Neumann, 35, 37, 184

Boundary-Value Problem, 12, 19, 61, 109, 183, 189

Box-Integration Discretization, 61

Brunetti, Rossella, *see* Band-Structure, Brunetti's

C

CAD (Computer-Aided Design), 2
- Electronic CAD, *see* ECAD
- hierarchy of, 2
- Technology CAD, *see* TCAD

Cartesian Coordinates System, *see* Wave Vector

CFD (Computational Fluid Dynamics), 6, 15
- diffusion equation, 35

Channel Length, 1, 135, 146f, *see also* Gate Length
- effect on MOSFET operation, **143**
- effective, 138, 145t
- future, 2

  gate, 133, 135, 145t, 146f
  metallurgical, 135, 138, 146f, 149f
Charge Neutrality, 36, 190
Chynoweth Model, 113
Circuit Simulation, 2
Classical Mechanics, 14
Code
  BJT simulator, 85
  MOSFET simulator, 137
  n^+nn^+simulator, 85
  validation by thermal-equilibrium, 9, 90
Collision, *see* Scattering
Computational Fluid Dynamics, *see* CFD
Computer Used, 85, 110, 137
Computer-Aided Design, *see* CAD
Conductivity, 63, 179
  at interface, 66, 181
Conservation
  in Boltzmann Transport Equation, 14
Contact
  ohmic, 35, **35**, 87, 89, 109, 111, 112, 134, 135, 184, 190
  ohmic, physics, 35
  Schottky, 35n
Control-Volume Discretization, *see* Discretization, control-volume
Convergence
  Gummel loop, 90, 103f
  matrix solvers, 90, 104f
  numerical solution, 90
  self-consistent solution, 90, 103f
  SOR, 90
Convergence Criterion
  Boltzmann equation, 84, 85, 136
  hole-continuity equation, 85
  Poisson equation, 85
Coordinates System, *see* Wave Vector
Current-Conservation Property
  of BTE, 14, 61, 112, **173**, 175
  of control-volume discretization, 61, 62, 88, 97f, 112, 122f, 140, 152
Curse of Dimensionality, 11, 14

D

Dark-Space Effect, 88, 112, 142, *see also* Dead-Space Effect
Dead-Space Effect
  in BJT, 112
  in MOSFET, 142
  in n^+nn^+, 88
Debye Length, 30, 49
Degenerate Equation, *see* Equation, degenerate elliptic
Density of States, *see* DOS
Device Simulation
  choice of variables in, 54
  closed-form solution, 3
  comparison of approaches, 7f
  definition, 3
  drift-diffusion, *see* Drift-Diffusion
  energy-transport, 4, 107
  hierarchy, 3
  hydrodynamic, 4, 7f, 54, 107, 141
  hydrodynamic as pre-processor, 6
  in TCAD hierarchy, 2, 3f
  Monte Carlo, *see* Monte Carlo
  numerical solution, 3
  spherical harmonic, *see* Spherical Harmonic
Diagonal Dominance
  advantage of, 73
  at low energy, 53
  Boltzmann equation, 53
  definition, 73

  - first derivatives and, 57n
  - hole-continuity equation, 192
  - impact ionization, 75, 76
  - inelastic acoustic phonon, 75, 76
  - inter-band optical phonon, 75, 76
  - optical phonon, 53, 74, 76
  - original variable, 74
  - Poisson equation, 187
  - pre-Maxwellian variable, 71, 73, 76, 78
  - second derivatives, 57n
  - space-dependent term, 74, 76
- Diffusion Equation, 35
- Dimensionality
  - curse of, 11, 14
  - of Boltzmann transport equation, 6, 14, 19
  - of distribution function, 6
  - reduction by spherical harmonics, 6, 14
- Dirichlet, *see* Boundary Condition, Dirichlet
- Dirichlet Boundary Condition, *see* Boundary Condition, Dirichlet
- Discretization
  - arithmetic-mean, 66, 180
  - Boltzmann equation, **61**, **63**
  - Boltzmann equation, final form, 68
  - boundary condition, 69
  - box-integration, 61
  - central-difference, 66, 185, 191
  - control-volume, 8, 61, 62, 136, 140, 175, 184
    - current-conserving property, 61, 88, 97f, 112, 122f, 140, 152f
  - control-volume face, 62, 63
  - control-volume principles, **62**
  - current-conserving, 61, 62, 175
  - discretization error, 15, 63, 66, 67, 191
  - energy boundary condition, 70
  - equations in matrix form, 69
  - finite-difference, 6, 15
  - flux conservation & optical phonon, 62
  - harmonic-mean, *see* Harmonic-Mean Scheme
  - hybrid, 57n
  - impact ionization, **68**
  - inelastic acoustic phonon, 8, **68**, 75
  - insulator boundary condition, 70
  - inter-band optical phonon, **67**
  - Liang-Goldsman-Mayergoyz scheme, 179, 182
  - ohmic boundary condition, 70
  - optical phonon, **67**, 74
  - scattering terms, **67**
  - Scharfetter-Gummel, *see* Scharfetter-Gummel Discretization
  - second-order accurate, 63, 66, 67, 185, 186, 191
  - space-dependent term, 74, 76
  - upwind, 57n
- Dispersion Relation, 15, **17**, 21, 24, **44**, 59
- Dissertation
  - organization, 9
  - summary & contribution, 8
- Distribution Function
  - asymmetric part, 19
  - definition of, 12
  - dimensionality, *see* Dimensionality
  - effect of impact ionization on, 113, 125f
  - energy distribution function, 19
  - equilibrium, *see* Thermal Equi-

librium
  exponential variation in energy of, 71, 73
  exponential variation in space of, 71, 73
  exponential variation of, 8
  impact ionization dependence on, 142
  in ballistic transport, 114, 142, 165f
  in BJT, 110, 119f, 120f
  in MOSFET, 140, 154f, 155f, 156f
  in n^+nn^+, 89, 90, 100f, 101f
  isotropic part of, 19
  Maxwellian, *see* Maxwellian Distribution
  non-Maxwellian, 141
  numerical noise, *see* Numerical Noise
  symmetric nature of, 6, 14
  symmetric part, 19
  tail, 5, 113, 114, 142
  thermal equilibrium, *see* Thermal Equilibrium
  total, **48**
  two populations, 89, 111, 141
DOS (Density of States), 23, 44, 47f
Drift-Diffusion, 4, 141
  as initial guess, 83
  as pre-processor, 6, 9, 136, 138, 147f
  BJT, 107, 109
  choice of variables in, 54
  comparison with BTE, 89, 114
  comparison with other approaches, 7f
  validity of, 4, 133

E

ECAD (Electronic CAD), 2, 3f
Effective Mass, 17, 44
Electron Concentration, 22
  in BJT, 111, 121f
  in higher bands, 142, 164f
  in MOSFET, 141, 142, 157f, 163f, 164f
  in n^+nn^+, 88, 96f
  potential-dependence, 73, 79, 184, 186
  Slotboom variables, 54, 72, 79, 184
Electron Current
  BJT, 112, 122f
  definition, 24
  n^+nn^+, 88, 97f
Electron Current Density
  MOSFET, 141, 158f
Electron Energy, 14
  in BJT, 112, 123f
  in n^+nn^+, 88, 98f
  MOSFET, 141, 160f, 162f, 168f
Electron Mass, *see* Effective Mass
Electron Spin, *see* Spin
Electron Velocity, *see* Velocity
Elliptic Equation, *see* Equation, elliptic
Energy-Transport, *see* Device Simulation, energy transport
Equation
  coupled differential, 83
  degenerate elliptic, 39
  differential, 19
  differential-difference, 14
  diffusion equation, 35
  dominant variable, 83
  elliptic, 35, 39
  integro-differential, 11, 14, 19
  ordinary differential, 191
  self-adjoint, 35
Equipment Simulation, 2

F

Feature Size, 1

Fermi's Golden Rule, 27
Finite-Difference Discretization, 6, 15
First Derivatives, disadvantage of, 57n
FORTRAN 77, 85, 137
  Gnu compiler, 85, 138

G

Gate Length, 133, 135, 146f, *see also* Channel Length
Gate Oxide, 36, 37n, 134, 137, 141, 146f
  thickness, 135, 145t, 146f
Gnudi, Antonio, 133
Goldsman, Prof. Neil, 6, 8, 179, 182
Grid
  BJT, 109, 110
  computational molecule, 65f, 180f, 185f, 190f
  for Boltzmann equation, 62, 63, 64f, 65f, 70f, 72f
  in Hamiltonian, 62
  in Hamiltonian and inter-band optical phonon, 62
  in Hamiltonian and optical phonon, 62
  in self-consistent solution, 90
  in space, 62
  MOSFET, **138**, 148f
  MOSFET grid generation, **136**
  near oxide in MOSFET, 137
  n^+nn^+, 87
  stencil, 65f
Gummel Loop, 83, 84, 86f, 87, 90
Gummel's Decoupled Scheme, 81, 83, 86f, 109, 186

H

Hamiltonian, 15, **17**, 19, 26, 54
Harmonic-Mean Scheme, 66, **179**, 181
  energy boundary condition and, 66
  interface conductivity, 66, 181
  similarity to Liang-Goldsman-Mayergoyz scheme, 179, 182
  similarity to Scharfetter-Gummel discretization, 179, **181**
HBT (Hetero-Junction BJT), 108
Heat Equation, 63, 179
  conductivity, 63
Hennacy, Kenneth A., 15
Hole Concentration
  in BJT, 112, 122f
  Slotboom variables, 184
Hole-Continuity Equation, 81, 181, **189**
  base boundary condition, 109
  boundary conditions, **190**
  in 1-D BJT, 109
  in BJT, 107
  in spherical-harmonic approach, 6
Hot Electrons, 73
  definition, 5
  distribution function, 89, 111, 141
  in BJT, 111
  in MOSFET, 141
Hot-Electron Effect
  in BJT, 111
  in MOSFET, 5, 142
  MOSFET substrate current, 139
Hydrodynamic, *see* Device Simulation, hydrodynamic

I

Impact Ionization
  BJT, 111, 112, 113, 124f
  Chynoweth model, 113
  effect on distribution function, 113, 125f

    generation due to, 173
    MOSFET, 141, 142, 161f, 163f
    MOSFET substrate current, 139
    n^+nn^+, 88, 98f
    threshold energy, 75
Impact Ionization Coefficient, **88**
    BJT, 112, 124f
    n^+nn^+, 88, 98f
Impact Ionization Scattering, 11, 27, **31**, 33f
    discretization, *see* Discretization
    energy-conservation in, 32
    pre-factor (P_{impact}), 50
    pre-Maxwellian variable, *see* Pre-Maxwellian Variable
    scattering rate, 50, 52f
    Thoma's expression, 34
    transport model in, 50
Integration Over k-Space, 20
Inter-Band Optical Phonon, 27, **28**, 62, 79, 111, 142, 174, 175
    discretization, *see* Discretization
    effect on upper-band population, 142
    pre-Maxwellian variable, *see* Pre-Maxwellian Variable
    scattering rate, 28
    similarity to optical phonon, 28
    transport model in, 49
Interface
    artificial, 36, 135
    semiconductor-insulator, 36
    semiconductor-metal, 35
    semiconductor-oxide, 36, 135
Intrinsic Concentration (n_i), 178, 184
Intrinsic Region of Device, 35
Ionized Impurity Scattering, 27, **30**
    Debye length, 30, 49
    discretization, *see* Discretization
    effect on velocity, 113, 126f
    elastic nature of, 30
    enhancement factor (ξ_{impurity}), 50
    equilibrium electron density in, 30
    pre-Maxwellian variable, *see* Pre-Maxwellian Variable
    scattering rate, 30, 50, 51f
    transition rate, 30
    transport model in, 49
Italian Spherical-Harmonic Group, 6, 15

J

Junction Depth in MOSFET, 134, 145t, 146f
Junction Isolation in BJT, 117f

K

k-Space
    conversion of sum to integral, 20
    integration over, 20
    summation over, 19
Kolodziejczak, K., 29

L

Law of Mass Action, 36, 190
Legendre Polynomials, 6, 30
Liang-Goldsman-Mayergoyz Scheme, 179, 182
Lin, Hongchin, 37

M

Maryland, University of, 6
Mass, *see* Effective Mass
Maxwell-Boltzmann Distribution, *see* Maxwellian Distribution
Maxwellian Distribution, 14, 36, 73, 78, 79, 88, 89, 90, 113, 114, 143, 177, 178

    as initial guess, 83, 136
    average energy of, 88, 115, 143
    BJT, 111, 115
    MOSFET, 140, 142
    n^+nn^+, 90, 101f
Maxwellian Factor, 54
Mayergoyz, Prof. Isaak, 6, 179, 182
    fixed-point technique advantages, 80
    fixed-point technique of, 8, 54, 72, 80
Meteorology, Spherical Harmonics in, 15
MIT (Massachusetts Institute of Technology), 6
Momentum
    quantum mechanical, 14
Monte Carlo, 4, 11, 80, 133, 141, 184
    BJT, 112
    BJT doping, 110
    BJT simulation, 108
    comparison with other approaches, 7f
    CPU-time comparison with spherical harmonic, 5, 7f
    full-band, 8
    retarding potential, 108
    self-consistent, 80, 184
    statistical noise in, 5
Moore's Law, 1
MOSFET, 133, **134**
    as 2-D device, 135
    ballistic transport, 142, 165f
    band-structure in simulation, 133, 134
    boundary conditions, **135**
    channel length, 1, 135, 146f
    channel length, effective, 138, 145t
    channel length, future, 2
    channel length, gate, 133, 135
    channel length, metallurgical, 135, 145t, 146f, 149f
convergence criterion, 136
current through cross-section, 140, 152f
DD versus BTE, 141, 157f
dead-space effect, 142
decoupled solution scheme, **136**, 147f
deep sub-micron, 133
distribution function, 140, 154f, 155f, 156f
distribution function at thermal equilibrium, 166f
doping profile, 138, 149f
effect of channel-length, **143**
electric field, 140, 141, 143, 153f, 162f, 168f
electric potential, 140, 141, 152f, 162f
electron concentration, 141, 157f, 163f
electron concentration in higher bands, 142, 164f
electron concentration spreading near drain, 141, 157f
electron current density, 141, 158f
electron energy, 141, 160f, 162f, 168f
electron velocity, 141, 142, 143, 159f, 163f, 168f
gate current, 5
grid, **138**, 148f
grid generation, **136**
grid near oxide, 137
half-Gaussian doping, 138
hot-electron, 139, 141, 142
hot-electron effects, *see* Hot-Electron Effects, in MOSFET
hot-electrons, 142
I_{DS}-V_{DS} characteristic, 139, 151

I_{DS}-V_{GS} characteristic, 139, 150f
impact ionization, 141, 142, 161f, 163f
I_{sub}-V_{GS} characteristic, 139, 151f
I-V characteristics, **139**, 150f, 151f
junction depth, 134, 145t, 146f
long channel-length, 143
output characteristics, 139, 151f
poly-silicon gate, 138
principle of operation, 134
saturation, 139
self-consistent simulation, 133
simulation time, 138
simulator, 137
spherical-harmonic simulation, 6, 9, 133
strong inversion, 139
structure, 134, 146f
structure for simulation, 138, 145t
sub-micron, 133, 134
sub-threshold, 139
sub-threshold transfer characteristic, 139, 150f
substrate current, 5, 139, 151f
substrate current & impact ionization, 139, 141
thermal equilibrium, **143**
thermal equilibrium electric field, 143, 167f
thermal equilibrium energy, 143, 167f
threshold voltage, 139, 145t, 150f
transconductance, 144, 169f
transfer characteristic, 139, 150f
transistor count, 1, 133
velocity, 141, 142, 143, 159f, 163f, 168f
velocity overshoot, 142, 143

N

Nabla Operator, 14
Neumann, *see* Boundary Condition, Neumann
n^+nn^+, 81, 85
    comparison with MOSFET, 81, 85
    DD versus BTE, 89, 99f
    dead-space effect, 88
    distribution function, 89, 90, 100f, 101f
    doping & structure, 87, 94f
    electric field, 87, 95f
    electric potential, 87, 95f
    electron concentration, 88, 96f
    electron current, 88, 97f
    electron energy, 88, 98f
    grid, 87
    Gummel loops, 87
    impact ionization coefficient, 88, 98f
    self-consistency, 88, 99f
    simulation time, 87
    simulator, 85
    spherical-harmonic simulation, 6, 9
    thermal equilibrium, 90, 101f
    thermal equilibrium electric field, 90, 102f
    velocity, 87, 96f
    velocity overshoot, 87
Non-Degenerate Semiconductors, 26
Non-Equilibrium Transport, 5
    BJT, 112
    MOSFET, 144
    n^+nn^+, 81, 85, 88
Non-Local Effect, 5
    BJT, 112
Non-Parabolicity Factor, 44

Numerical Noise, 8, 81, **91**
  due to absence of coupling, 91
  first derivatives, 57
  inelastic acoustic phonons and, **91**, 93, 105f, 106f
  numerical cause, 92
  physical cause, 92

O

Optical Phonon, 27, **27**, 79
  Bose-Einstein relation in, 27
  conservation of flux, 62
  deformation potential, 27
  discretization, *see* Discretization
  energy exchange in, 93
  energy of, 27
  frequency of, 27
  generation due to, 174, 175
  number of, 27
  pre-Maxwellian variable, *see* Pre-Maxwellian Variable
  scattering rate, 28, 50f
  transition rate, 27
  transport model in, 48
  velocity saturation, 114
Orthogonality, *see* Spherical Harmonic, orthogonality
Oxide, *see* Gate Oxide

P

p-n Junction, 107
Patankar, Suhas V., 66, 71, 179
  harmonic-mean scheme of, *see* Harmonic-Mean Scheme
Permittivity of free space, 183
Permittivity of Silicon, 30, 183
Phonon
  acoustic, *see* Acoustic Phonon
  inter-band optical, *see* Inter-Band Optical Phonon
  optical, *see* Optical Phonon
Planck's Constant, 14, 44
Poisson Equation, 2, 183, **183**
  base boundary condition, 109
  Boltzmann-Poisson system, 9, 177, 178, 186
  boundary conditions, **184**
  damping, 186
  discretization, **184**
  in 1-D BJT, 109
  linear, 183, 184
  Mayergoyz' fixed-point technique, *see* Mayergoyz, fixed-point technique of
  non-linear, 72, 80, 90, 177, 178, 183, 184
  self-consistent, 6
  solvers, 72
  thermal equilibrium, *see* Thermal Equilibrium
Poly-Silicon, 138
Pre-Maxwellian Variable, 8, **54**
  acoustic phonon, **57**, 78t
  advantages of, 8, 53, **71**
  Boltzmann equation in, 55, **60**
  code reuse, 72
  diagonal dominance, 71, 73, 76, 78, 78t
  impact ionization, **59**, 78t, 79
  inter-band optical phonon, **57**, 78t
  ionized impurity scattering, **59**, 78t
  linearity of BTE, 8, 54, 72
  Maxwellian factor, 54
  optical phonon, **56**, 78t
  Poisson equation, 72, 79
  Slotboom variables, 54, 72, 79
  space-dependent factor, 54
Process Simulation, 2

Q

Quantum Mechanics, 12, 14
Quasi-Fermi Potential Variables, 54

R

Reduced Planck's Constant, 14
Reliability, 8
  MOSFET, 142

S

Scattering, **26**
  acoustic, *see* Acoustic Phonon
  analytical evaluation of, 6
  Collision, **26**
  discretization, *see* Discretization
  energy exchange in, 92
  impact ionization, *see* Impact Ionization Scattering
  in non-degenerate semiconductors, 26
  inter-band optical, *see* Inter-Band Optical Phonon
  ionized impurity, *see* Ionized Impurity Scattering
  optical, *see* Optical Phonon
  phonon, *see* Phonon
  pre-Maxwellian variable, 56
  randomizing nature of, 6, 14
  spin in, 26
  total rate, 35
  transition rate, *see* Transition Rate
  transport model, *see* Transport Model
Scattering Integral
  analytical evaluation of, 6
Scattering-Matrix Approach, 11, 108
Scharfetter-Gummel Discretization, 57n, **190**
  similarity to harmonic-mean scheme, 179, **181**
Schrödinger Equation, 19
Second Derivatives, advantage of, 57n
Seidman Modification, 192
Self Adjoint
  Boltzmann equation, 18
  inelastic acoustic phonon, 8, 9
Semiconductors
  non-degenerate, 26
  rice of industry, 1
Silicon
  density of, 27
  DOS (density of states), *see* DOS
  permittivity of, 30, 183
  relative permittivity of, 183
  sound velocity in, 29
Silicon Dioxide, 135
Simulation
  BJT, *see* BJT
  device, *see* Device Simulation
  equipment, 2
  MOSFET, *see* MOSFET
  n^+nn^+, *see* n^+nn^+
  process, 2, 54
  small-circuit, 2
Simulator
  BJT, 85
  MOSFET, 137
  n^+nn^+, 85
Slotboom Variables, 54, 72, 79, 80, 179, 184
SOR (Successive Over-Relaxation), 8, 81, 90
  line, **82**, 83, 84f, 90, 91, 136
  point, 90, 91
  rate of convergence, 91
  SOR factor, 83, 90
Spectral Methods
  errors in, 15, 18
  spherical harmonic as, *see* Spherical Harmonic, as spectral method
Spherical Coordinates System, *see* Wave Vector

Spherical Harmonic
  advantages, 5
  as spectral method, 6, 15, 18
  band-structure in, 8
  basic idea, 5
  BJT, *see* BJT
  Boltzmann equation as degenerate elliptic, 39
  comparison with other approaches, 5, 7f
  CPU-time comparison with Monte Carlo, 5, 7f
  expanded Boltzmann equation, 18, 19
  full-band band-structure in, 8
  functional form of, 16
  hybrid simulator, 6
  impact ionization studied in, 6
  in Boltzmann transport equation
    summary of derivation, 15
  in meteorology, 15
  justification for first-order, 18
  literature review, 6
  MOSFET, *see* MOSFET
  n^+nn^+, *see* n^+nn^+
  orthogonality of, 15, 16, 23, 25
  physical significance of coefficients, 19
  post-processing, 6, 9
  second-order differential equation, 19
  self-consistent, 6, 9, 108, 133, 186
  space-dependent, 8
  space-independent, 6, 8
  summary of derivation, 15
  truncation, 15, 18
Spin
  in scattering, 26
  in summation over k-space, 20
Substrate Current, *see* MOSFET, substrate current
Summation Over k-Space, 19

T

TCAD (Technology CAD), 2, 3f
Thermal Diffusion, 141
Thermal Equilibrium, 9, **89**, **177**
  analytical solution at, 9, 89, 177
  BJT, 115
  Boltzmann-Poisson system, 177, 178
  code validation, *see* Code, validation by thermal-equilibrium
  definition of, 177
  distribution function at, 6, 14, 89, 90, 177, 178
  distribution function in BJT, 115, 130f
  distribution function in MOSFET, 166f
  distribution function in n^+nn^+, 90, 101f
  electric field in BJT, 116, 131f
  electric field in MOSFET, 143, 167f
  electric field in n^+nn^+, 90, 102f
  electron energy in BJT, 115, 131f
  electron velocity at, 9
  electron velocity in MOSFET, 143
  energy in MOSFET, 143, 167f
  MOSFET, **143**
  motivation to study, 89
  n^+nn^+, 90, 101f
  unique solution at, 178
Thermal Voltage, 178, 186, 189
Thesis, *see* Dissertation
Thoma, R., 34, 50
Thomas' Tridiagonal Matrix Algorithm, 83, 187, 192

Threshold Voltage, 139, 145t, 150f, *see also* MOSFET, threshold voltage
  linear extrapolation method, 139
Transconductance, *see* MOSFET, transconductance
Transistor Count, 1
Transition Rate, 26
  for acoustic phonon scattering, 29
  for ionized impurity scattering, 30
  for optical phonon scattering, 27
Transport Model, **48**
  acoustic phonon, 49
  BJT, 48
  bulk Silicon, 48
  impact ionization scattering, 50
  inter-band optical phonon, 49
  ionized impurity scattering, 49
  low-field, 48
  MOSFET, 48
  optical phonon, 48
Tridiagonal Matrix Algorithm, 83, 187, 192

U

University of Bologna, 6
University of Maryland, 6

V

Vecchi, Maria C., 8, 59
Velocity
  average velocity definition, 24
  BJT, 111, 121f
  damping, 9, 89, 116, 143
  derivations, 23
  effect of ionized impurity scattering, 113, 126f
  graded-base BJT, 115, 129f
  group velocity, 12, 23, 24
  MOSFET, 141, 142, 143, 159f, 163f, 168f
  n^+nn^+, 87, 96f
  overshoot
    BJT, 111, 114
    cause, 87
    definition, 87n
    graded-base BJT, 115
    MOSFET, 142, 143
    n^+nn^+, 87, 89
  overshoot & ionized impurity scattering, 113, 114
  saturation, 87, 88, 114
  undershoot, 9, 89, 116, 143
Ventura, Davide, 133

W

Wave Vector
  Cartesian coordinates, 15, 16f
  differential element in spherical coordinates, 20
  physical significance, 15
  spherical coordinate, 15, 16f
Weather Prediction, Spherical Harmonics in, 15
White, Prof. Jacob, 6

Z

Zero Neumann, *see* Boundary Condition, zero Neumann