Hadis Morkoç

Handbook of Nitride Semiconductors and Devices

Related Titles

Piprek, J. (ed.)

Nitride Semiconductor Devices: Principles and Simulation

2007

ISBN: 978-3-527-40667-8

Adachi, S.

Properties of Group-IV, III-V and II-VI Semiconductors

2005

ISBN: 978-0-470-09032-9

Ruterana, P., Albrecht, M., Neugebauer, J. (eds.)

Nitride Semiconductors

Handbook on Materials and Devices

2003

ISBN: 978-3-527-40387-5

Ng, K. K.

Complete Guide to Semiconductor Devices

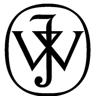
2002

ISBN: 978-0-471-20240-0

Hadis Morkoç

Handbook of Nitride Semiconductors and Devices

Vol. 2: Materials Properties, Physics and Growth



WILEY-VCH Verlag GmbH & Co. KGaA

The Author

Prof. Dr. Hadis Morkoç
Virginia Commonwealth University
Dept. of Chemical Engineering
Richmond, VA
USA

Cover
SPIESZDESIGN GbR,
Neu-Ulm, Germany

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Typesetting Thomson Digital, Noida, India
Printing Strauss Gmbh, Mörlenbach
Binding Litges & Dopf GmbH, Heppenheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN: 978-3-527-40838-2

Contents

Preface *XIII*

Color Tables *XXI*

1 Metal Contacts to GaN and Processing *1*
Introduction *1*
1.1 A Primer for Semiconductor–Metal Contacts *2*
1.2 Current Flow in Metal–Semiconductor Junctions *6*
1.2.1 The Regime Dominated by Thermionic Emission *10*
1.2.2 Thermionic Field Emission Regime *12*
1.2.3 Direct Tunneling Regime *20*
1.2.4 Leakage Current *24*
1.3 GaN Schottky Barriers for High-Voltage Rectifiers *33*
1.4 Ohmic Contact Resistance *38*
1.4.1 Specific Contact Resistivity *38*
1.4.2 Semiconductor Resistance *39*
1.5 Determination of the Contact Resistivity *40*
1.6 Ohmic Contacts to GaN *46*
1.6.1 Nonalloyed Ohmic Contacts *46*
1.6.2 Alloyed Ohmic Contacts on n-Type GaN *47*
1.6.3 Contacts to p-Type GaN and Transparent Conducting Oxides *53*
1.6.4 Effect of Surface Treatment on Ohmic Contacts *56*
1.6.5 Case of a Forward-Biased p-n Junction in Conjunction with Nonohmic Contacts to p-GaN *58*
1.7 Structural Analysis of Ohmic Contacts on GaN *60*
1.8 Etching Techniques for III Nitrides *65*
1.8.1 Dry (Plasma) Etching *66*
1.8.1.1 Electron Cyclotron Resonance Etching *67*
1.8.1.2 Ion Milling *71*
1.8.1.3 Reactive Ion Etching *73*

Handbook of Nitride Semiconductors and Devices. Vol. 2. Hadis Morkoç
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40838-2

1.8.1.4 Inductively Coupled Plasma Etching 75
1.8.1.5 Selective Etching of GaN/AlGaN 77
1.8.1.6 Dry Etching of p-GaN 78
1.8.1.7 Dry Etching on Ga- and N-Face of Freestanding GaN Substrate 78
1.8.1.8 Magnetron Reactive Ion Etching 79
1.8.1.9 Chemically Assisted Ion Beam Etching (CAIBE) 80
1.8.1.10 RF Plasma Etching of GaN 81
1.8.2 Laser Ablation Etching of GaN 83
1.8.3 Wet Etching 84
1.8.4 Photochemical Etching 92
1.9 Implant Isolation 99
1.10 Process Damage 102
References 108

2 Determination of Impurity and Carrier Concentrations 121
Introduction 121
2.1 Impurity Binding Energy 121
2.2 Conductivity Type: Hot Probe and Hall Measurements 122
2.2.1 Measurement of Mobility 122
2.3 Semiconductor Statistics, Density of States, and Carrier Concentration 125
2.3.1 Degeneracy Factor 132
2.3.2 Charge Balance Equation and Carrier Concentration 135
2.3.2.1 n-Type Semiconductor 136
2.3.2.2 p-Type Semiconductor 142
2.3.2.3 Multiple Occupancy of the Valence Bands 146
2.4 Capacitance–Voltage Measurements 147
Appendix 2.A. Fermi Integral 158
Appendix 2.B. Density of States in 3D, 2D, and 1D Systems 159
2.B.1. Three-Dimensional Structure 159
2.B.2. Two-Dimensional Structure 160
2.B.3. One-Dimensional Structure 161
References 162

3 Carrier Transport 165
3.1 Prelude 165
3.2 Carrier Scattering 169
3.2.1 Boltzmann Transport Equation 171
3.2.2 Impurity Scattering 182
3.2.2.1 Ionized Impurity Scattering 183
3.2.2.2 Neutral Impurity Scattering 190
3.2.3 Acoustic Phonon Scattering 192
3.2.3.1 Deformation Potential Scattering 194
3.2.3.2 Piezoelectric Scattering 198
3.2.4 Optical Phonon Scattering 201

3.2.4.1 Nonpolar Optical Phonon Scattering *201*
3.2.4.2 Polar Optical Phonon Scattering *203*
3.2.5 Short-Range Potential-Induced Scattering *210*
3.2.5.1 Alloy Potential-Induced Scattering *212*
3.2.5.2 Potential Barrier Scattering *216*
3.2.5.3 Potential Well Scattering *218*
3.2.5.4 Space Charge Scattering *218*
3.2.5.5 Dipole Scattering *219*
3.2.6 Carrier–Carrier Scattering *222*
3.2.7 Plasmon Scattering *223*
3.2.8 Boundary Scattering *225*
3.2.9 Dislocation Scattering *225*
3.2.10 Effect of Inhomogeneities on Mobility *231*
3.3 Calculated Mobility of GaN *232*
3.4 Scattering at High Fields *239*
3.4.1 Transport at High Fields: Energy and Momentum Relaxation Times *247*
3.4.2 Electron–Phonon Interaction *248*
3.4.2.1 Polar Optical Phonon Scattering *250*
3.4.2.2 Nonpolar Optical Phonon Scattering *251*
3.4.2.3 Acoustic Phonon (Deformation Potential) Scattering *251*
3.4.2.4 Acoustic Phonon Piezoelectric Scattering *251*
3.4.3 High-Field Transport in Single- and Multivalley Models *252*
3.5 Measurements of Mobility and Associated Fundamentals *261*
3.5.1 Magnetoresistance *261*
3.5.2 Hall Effect and Physical Magnetoresistance *268*
3.5.2.1 First-Order Mobility Measurement: A Textbook Case *270*
3.5.2.2 Higher Order Mobility Measurements with Energy-Dependent Relaxation Time and Hall Factor *274*
3.5.2.2.1 Energy-Independent Relaxation Time and Arbitrary B *275*
3.5.2.2.2 Energy-Dependent Relaxation Time *276*
3.5.3 Geometric Magnetoresistance (GMR) *281*
3.5.3.1 Energy-Dependent Relaxation Time $B \to 0$ and $\omega_c\tau$ *281*
3.5.3.2 Energy-Dependent Relaxation Time and High B Field, $\omega_c\tau \gg 1$ *282*
3.5.4 Hall Factor *282*
3.6 Multiband Effects and Mixed Conductivity *283*
3.6.1 Energy-Dependent Relaxation Time with Low B field ($\omega_c\tau \gg 1$) *286*
3.6.2 Energy-Dependent Relaxation Time and High B Field ($\omega_c\tau \ll 1$ for both carriers) *286*
3.6.3 Energy-Dependent Relaxation Time and High B Field, $B \to \infty$ *287*
3.6.4 Quantitative Mobility Spectrum Analysis *289*
3.7 van der Pauw Method *295*
3.8 Quantum Hall Effect – Shubnikov–de Haas *301*
3.9 Measured Mobility in n-type GaN *315*
3.9.1 Mobility in GaN *315*

3.9.2 Delineation of Multiple Conduction Layer Mobilities *326*
3.9.3 Transport Along the *c*-Direction *329*
3.10 Measurement of High-Field Electron Velocity in n-Type GaN *330*
3.11 Carrier Transport in p-Type GaN *333*
3.12 Carrier Transport in InN *335*
3.13 Carrier Transport in AlN *339*
3.14 Transport in Unintentionally Doped and High-Resistivity GaN *341*
3.15 Carrier Transport in Alloys *343*
3.16 Two-Dimensional Transport in n-Type GaN *348*
3.16.1 General Scattering Theory in 2D Systems *351*
3.16.1.1 Ionized Impurity Scattering in 2D Systems *356*
3.16.1.2 Acoustic Phonon Deformation Potential Scattering *358*
3.16.1.3 Acoustic Phonon Piezoelectric Potential Scattering *359*
3.16.1.4 Polar Optical Phonon Scattering *360*
3.16.2 A General Discussion of Electron Mobility in AlGaN/GaN 2D System *361*
3.16.3 Low-Field Transport in 2DEG Systems *367*
3.16.3.1 Low Temperature *368*
3.16.3.2 High Temperature *369*
3.16.3.3 Interface Phonon Scattering *369*
3.16.3.4 Fröhlich Scattering in Two-Dimensional Electron Gas *369*
3.16.4 Numerical Two-Dimensional Electron Gas Mobility Calculations *370*
3.16.5 Two-Dimensional Hole Gas *378*
3.17 Interface Roughness Scattering *385*
3.18 Quantum Transport in AlGaN/GaN 2DEG *388*
3.19 Observations *394*
References *395*

4 The p–n Junction *409*
Introduction *409*
4.1 Heterojunctions *409*
4.2 Band Discontinuities *410*
4.2.1 GaN/AlN Heterostructures *414*
4.2.2 GaN/InN and AlN/InN *416*
4.3 Electrostatic Characteristics of p–n Heterojunctions *428*
4.4 Current–Voltage Characteristics of p–n Junctions *435*
4.4.1 Generation–Recombination Current *436*
4.4.2 Surface Effects *440*
4.4.3 Diode Current Under Reverse Bias *442*
4.4.4 Field Effects and Hopping *442*
4.4.4.1 Poole–Frenkel Current *443*
4.4.4.2 Hopping Current *449*
4.4.5 Avalanche Multiplication *454*
4.4.6 Diffusion Current *459*
4.4.7 Diode Current Under Forward Bias *463*

4.5 *I–V* Characteristics of Ideal GaN-Based p–n Junctions *464*
4.6 *I–V* Characteristics of GaN-Based p–n Junctions *468*
4.6.1 GaN p–n Junctions *468*
4.6.1.1 Forward Bias *I–V* *468*
4.6.1.2 Reverse Bias *I–V* *474*
4.6.2 AlGaN–GaN Heterojunctions *480*
4.6.3 AlGaN Heterojunctions *481*
4.7 High-Voltage Blocking GaN and AlGaN-Based p–n Junctions *483*
References *485*

5 Optical Processes in Semiconductors and Optical Properties of Nitride Semiconductors and Heterostructures *491*
Introduction *491*
5.1 Basics of Photoluminescence *498*
5.1.1 Steady-State Photoluminescence *499*
5.1.2 Recombination Statistics *500*
5.1.3 Effect of Temperature on PL Intensity *501*
5.1.4 Quantum Efficiency *503*
5.1.5 Effect of Excitation Intensity on PL Intensity *504*
5.1.6 Determination of Acceptor Concentration in n-type GaN *504*
5.1.7 Time-Resolved Luminescence *505*
5.1.8 Photoluminescence Excitation Spectra *507*
5.1.9 Spatially and Depth-Resolved Cathodoluminescence *508*
5.1.10 Optically Detected Magnetic Resonance *509*
5.2 Band-to-Band Transitions *509*
5.2.1 Excitonic Transitions *511*
5.3 Optical Transitions in GaN *512*
5.3.1 Excitonic Transitions in GaN *513*
5.3.1.1 Free Excitons *514*
5.3.1.1.1 Polaritons *517*
5.3.1.2 Bound Excitons *542*
5.3.1.2.1 Excitons Bound to Shallow Donors *547*
5.3.1.2.2 Excitons Bound to Acceptors *560*
5.3.1.3 Excitons in Magnetic Field *562*
5.3.2 Exciton Recombination Dynamics *575*
5.3.2.1 Exciton Recombination Dynamics at Low-to-Medium Injection Levels *578*
5.3.2.2 Recombination Dynamics at High Injection Levels *587*
5.3.3 Free-to-Bound Transitions *589*
5.3.4 Donor–Acceptor Transitions *591*
5.3.5 Luminescence Related to Structural Defects *593*
5.3.5.1 Effect of Excitation Intensity *599*
5.3.5.2 Effect of Temperature *600*
5.3.5.3 Properties of the Yi Peaks *606*

5.3.6 Emission from Deep-Level Defects *610*
5.3.6.1 Yellow Luminescence *610*
5.4 Group-II Element Related Transitions *620*
5.4.1 Role of Potential Fluctuations in GaN *620*
5.4.2 Role of Coulombic Interaction in Transition Energies in GaN *623*
5.4.3 UVL and BL Bands in Compensated and Heavily Mg-Doped GaN *623*
5.4.4 Local Structure of the Blue Band in GaN:Mg *631*
5.4.5 Blue Luminescence Band in Zn-Doped GaN *632*
5.5 Blue Luminescence Band in Undoped GaN *633*
5.5.1 Characteristics of Blue Luminescence in Undoped GaN *634*
5.5.2 Comparison of Photoluminescence in Undoped and Zn-Doped GaN *636*
5.6 Surface-Related Blue Luminescence in Etched GaN *638*
5.7 Optical Properties of GaN Doped with Rare Earths *642*
5.7.1 Luminescence of Rare Earth in GaN *643*
5.7.1.1 Red Emission in Eu-Doped GaN *643*
5.7.1.2 Red Emission in Pr-Doped GaN *648*
5.7.1.3 Green and Infrared Emission in Er-Doped GaN *649*
5.7.1.4 Blue Emission in Tm-Doped GaN *652*
5.8 Optical Properties of Alloys *653*
5.8.1 Localization in Alloys *654*
5.8.2 Optical Properties of AlGaN *656*
5.8.3 Optical Properties of InGaN *661*
5.8.4 Optical Properties of InAlN *665*
5.9 Optical Properties of Nitride Heterostructures *666*
5.9.1 GaN/AlGaN Quantum Wells *668*
5.9.1.1 Polarization and Carrier Dynamics in GaN/AlGaN Quantum Wells *678*
5.9.1.1.1 Polarization in Quantum Wells with Implications to GaN/AlGaN Wells *679*
5.9.1.1.2 Optical Properties of Non-*c*-Plane GaN Based Structures *687*
5.9.1.1.3 Carrier Dynamics in GaN/AlGaN Quantum Wells *694*
5.9.2 Optical Properties of GaN/AlGaN Modulation-Doped Heterostructures *698*
5.9.3 InGaN/GaN, InGaN/InGaN, and InGaAlN/InGaAlN Quantum Wells *705*
5.9.3.1 InGaN/GaN, InGaN/InGaN Quantum Wells *707*
5.9.3.2 Localization and Polarization in InGaN Quantum Wells *708*
5.9.3.3 Size Dependence of the LO-Phonon Coupling in InGaN Quantum Wells *714*
5.9.3.4 InGaAlN/InGaAlN Quantum Wells *720*
5.10 Quantum Dots *723*
5.10.1 Effects of Quantum Confinement, Strain and Polarization *724*
5.10.2 GaN Quantum Dots *729*

5.10.3 InGaN Quantum Dots 740
5.11 Intraband or Intersubband Transitions in GaN/AlGaN Quantum Wells 745
5.12 Nonlinear Optical Properties of III-Nitrides 756
5.12.1 Second-Order Nonlinear Optical Properties 762
5.12.1.1 Second Harmonic Generation (SHG) 762
5.12.2 Third-Order Nonlinear Optical Properties 782
5.12.2.1 Intensity-Dependent Refractive Index 783
5.12.2.2 Two-Photon Absorption 787
5.12.2.3 Third Harmonic Generation (THG) 789
References 794

Index 829

Appendix 847

Preface

This three-volume handbook represents the only comprehensive treatise on semiconductor and device fundamentals and technology under the overall umbrella of wide bandgap nitride semiconductors with comparison to GaAs when applicable. As it stands, the book is a reference book, a handbook, and a graduate text book all in one and would be beneficial to second-year graduate students majoring in semiconductor materials and devices, graduate research assistants conducting research in wide bandgap semiconductors, researchers and technologists, faculty members, program monitors, and managers. The philosophy of this endeavor is to present the material as much clearly and comprehensively as possible, so that there is very little need to refer to other sources to get a grasp of the subjects covered. Extreme effort has been expended to ensure that concepts and problems are treated starting with their fundamental basis so that the reader is not left hanging in thin air. Even though the treatise deals with GaN and related materials, the concepts and methods discussed are applicable to any semiconductor.

The philosophy behind *Nitride Semiconductors and Devices* was to provide an adequate treatment of nitride semiconductors and devices as of 1997 to be quickly followed by a more complete treatment. As such, *Nitride Semiconductors and Devices* did not provide much of the background material for the reader and left many issues unanswered in part because they were not yet clear to the research community at that time. Since then, tremendous progress both in the science and engineering of nitrides and devices based on them has been made. While LEDs and lasers were progressing well even during the period when *Nitride Semiconductors and Devices* was written, tremendous progress has been made in FETs and detectors in addition to LEDs and lasers since then. LEDs went from display devices to illuminants for lighting of all kinds. Lasers are being implemented in the third generation of DVDs. The power amplifiers are producing several hundred watts of RF power per chip and the detectors and detector arrays operative in the solar-blind region of the spectrum have shown detectivities rivaling photomultiplier tubes. The bandgap of InN has been clarified which now stands near 0.7 eV. Nanostructures, which did not exist

Handbook of Nitride Semiconductors and Devices. Vol. 2. Hadis Morkoç
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40838-2

during the period covered by *Nitride Semiconductors and Devices*, have since become available. The technological breakthroughs such as epitaxial lateral overgrowth, laser liftoff, and freestanding GaN were either not fully developed or did not exist, neither did the highly improved quantum structures and devices based on them. In the interim period since then, the surfaces of nitrides and substrate materials, point defects and doping, magnetic ion doping, processing, current conduction mechanisms, and optical processes in bulk and quantum structures have been more clearly understood and many misconceptions (particularly, those dealing with polarization) identified, removed and/or elucidated. The handbook takes advantage of the fundamental and technological developments for a thorough treatment of all aspects of nitride semiconductors. In addition, the fundamentals of materials physics and device physics that are provided are applicable to other semiconductors, particularly, wurtzitic direct bandgap semiconductors.

The handbook presents a thorough treatment of the science, fundamentals, and technology of nitride semiconductors and devices in such a width and depth that the reader would seldom need to engage in time-consuming exploration of the literature to fill in gaps. Last but not the least, the handbook contains seamless treatments of fundamentals needed or relied on throughout the entire book. The following is a succinct odyssey through the content of the three-volume handbook.

Volume 1, Chapter 1 discusses the properties of nitride-based semiconductors with plenty of tables for reference. Volume 1, Chapter 2 treats the band structure of III–V nitrides, theories applied to determining the band structure, features of each theory with a succinct discussion of each, band structure of dilute III–V semiconductors doped with N, strain and stress, deformation potentials, and in-depth discussion of piezo and spontaneous polarization with illustrative and instructive artwork. Volume 1, Chapter 3 encompasses substrates that have been and are used for growth of nitride semiconductors, mainly, structural and mechanical (thermal) properties of those substrates, surface structure of planes used for growth, and substrate preparation for growth. Orientation and properties of GaN grown on those substrates are discussed along with commonly used surface orientations of GaN. The discussion is laced with highly illustrative and illuminating images showing orientations of GaN resulting through growth on *c*-plane, *a*-plane, *m*-plane, and *r*-plane substrates whichever applicable and the properties of resulting layers provided. The treatment segues into the discussion of various growth methods used for nitrides taking into account the fundamentals of growth including the applicable surface-oriented processes, kinetics, and so on, involved. A good deal of growth details for both OMVPE and MBE, particularly, the latter including the fundamentals of *in situ* process monitoring instrumentation such as RHEED, and dynamics of growth processes occurring at the surface of the growing layer are given. Of paramount interest is the epitaxial lateral overgrowth (ELO) for defect reduction. In addition to standard single multistep ELO, highly attractive nanonetwork meshes used for ELO are also discussed. Specifics in terms of growth of binary, ternary, and quaternaries of nitride semiconductors are discussed. Finally, the methods used to grow nanoscale structures are treated in sufficient detail.

Volume 1, Chapter 4 focuses on defects, both extended and point, doping for conductivity modulation and also for rendering the semiconductor potentially ferromagnetic segueing into electrical, optical, and magnetic properties resulting in films, with sufficient background physics provided to grasp the material. A clear discussion of extended defects, including line defects, are discussed with a plethora of illustrative schematics and TEM images for an easy comprehension by anyone with solid-state physics background. An in-depth and comprehensive treatment of the electrical nature of extended defects is provided for a full understanding of the scope and effect of extended defects in nitride semiconductors, the basics of which can be applicable to other hexagonal materials. The point defects such as vacancies, antisites, and complexes are then discussed along with a discussion of the effect of H. This gives way to the methods used to analyze point defects such as deep level transient spectroscopy, carrier lifetime as pertained to defects, positron annihilation, Fourier transform IR, electron paramagnetic resonance, and optical detection of magnetic resonance and their application to nitride semiconductors. This is followed by an extensive discussion of n-type and p-type doping in GaN and related materials and developments chronicled when applicable. An in-depth treatment of triumphs and challenges along with codoping and other methods employed for achieving enhanced doping and the applicable theory has been provided. In addition, localization effects caused by heavy p-type doping are discussed. This gives way to doping of, mainly, GaN with transition elements with a good deal of optical properties encompassing internal transition energies related to ion and perturbations caused by crystal field in wurtzitic symmetry. To get the reader conditioned for ferromagnetism, a sufficient discussion of magnetism, ferromagnetism, and measurement techniques (magnetic, magneto transport, magneto optics with underlying theory) applied to discern such properties are given. This is followed by an in-depth and often critical discussion of magnetic ion and rare earth-doped GaN, as well as of spintronics, often accompanied by examples for materials properties and devices from well-established ferromagnetic semiconductors such as Mn-doped GaN and Cr-doped ZnTe.

Volume 2, Chapter 1 treats metal semiconductor structures and fabrication methods used for nitride-based devices. Following a comprehensive discussion of current conduction mechanisms in metal semiconductor contacts, which are applicable to any metal semiconductor system, specific applications to metal-GaN contacts are treated along with the theoretical analysis. This gives way to a discussion of ohmic contacts, their technology, and their characterization. In particular, an ample discussion of the determination of ohmic contact resistivity is provided. Then etching methods, both dry (plasma) and wet, photochemical, process damage, and implant isolation are discussed. Volume 2, Chapter 2 deals with determination of impurity and carrier concentrations and mobility mainly by temperature-dependent electrical measurements, such as Hall measurements. Charge balance equations, capacitance voltage measurements, and their intricacies are treated and used for nitride semiconductors, as well as a good deal of discussion of often brushed off degeneracy factors.

Volume 2, Chapter 3 is perhaps one of the most comprehensive discussions of carrier transport in semiconductors with applications to GaN. After a discussion of scattering processes in physical and associated mathematical terms, the methods discussed are applied to GaN and other related binaries and ternaries with useful ranges of doping levels, compositions, and lattice temperatures. Comparisons with other semiconductors are also provided when applicable. This treatment segues into the discussion of carrier transport at high electric fields applicable to field-effect transistors, avalanche and pin (biased) photodiodes. This is followed by the measurement of mobility and associated details, which are often neglected in text and reference books. The discussion then flows into magnetotransport beyond that present in standard Hall measurements. Low, medium, and high magnetic field cases, albeit only normal to the surface of the epitaxial layers, are treated. The treatise also includes cases where the relaxation time, if applicable, is energy-dependent and somewhat energy-independent. The discussion of the magnetotransport paves the way for a fundamental and reasonably extensive discussion of the Hall factor for each of the scattering mechanisms that often is not treated properly or only in a cursory manner in many texts leading to confusion. After providing the necessary fundamentals, the transport properties of GaN are discussed. This gives way to the discussion of various scattering mechanisms in two-dimensional systems that are relied on in high-performance FETs. For determining the mobility of each layer (in the case of multiple layer conduction), quantitative mobility spectrum analysis including both the fundamentals and experimental data obtained in nitride semiconductors is discussed. The quantum Hall effect and fractional quantum Hall effect in general and as germane to GaN are discussed along with parameters such as the effective mass determined from such measurements.

Volume 2, Chapter 4 is devoted to p–n junctions, beginning with the discussion of band lineups, particularly, in the binary pairs from the point of view of theoretically and experimentally measured values. Current conduction mechanisms, such as diffusion, generation-recombination, surface recombination, Poole–Frenkel, and hopping conductivity are discussed with sufficient detail. Avalanche multiplication, pertinent to the high-field region of FETs, and avalanche photodiodes, are discussed-followed by discussions of the various homojunction and heterojunction diodes based on nitrides.

Volume 2, Chapter 5 is perhaps the most comprehensive discussion of optical processes that can occur in a direct bandgap semiconductor and, in particular, in nitride-based semiconductors and heterostructures inclusive of 3, 2-, and 0-dimensional structures as well as optical nonlinearities. Following a treatment of photoluminescence basics, the discussion is opened up to the treatment of excitons, exciton polaritons, selection rules, and magneto-optical measurements followed by extrinsic transitions because of dopants/impurities and/or defects with energies ranging from the yellow and to the blue wavelength of the visible spectrum. Optical transitions in rare earth-doped GaN, optical properties of alloys, and quantum wells are then discussed with a good deal of depth, including localization effects and their possible sources particularly media containing InN. The discussion then leads to the

treatment of optical properties of quantum dots, intersubband transitions in GaN-based heterostructures, and, finally, the nonlinear optical properties in terms of second and third harmonic generation with illuminating graphics.

Volume 3, Chapter 1 is devoted, in part, to the fundamentals of light emitting diodes, the perception of vision and color by human eye, methodologies used in conjunction with the chromaticity diagram and associated international standards in terms of color temperatures and color rendering index. Specific performances of various types of LEDs including UV varieties, current spreading or the lack of related specifics, analysis of the origin of transitions, and any effect of localization are discussed. A good deal of white light and lighting-related standards along with approaches employed by LED manufacturers to achieve white light for lighting applications is provided. The pertinent photon conversion schemes with sufficient specificity are also provided. Finally, the organic LEDs, as potential competitors for some applications of GaN-based LEDs are discussed in terms of fundamental processes that are in play and various approaches that are being explored for increased efficiency and operational lifetime.

Volume 3, Chapter 2 focuses on lasers along with sufficient theory behind laser operation given. Following the primer to lasers along with an ample treatment based on Einstein's A and B coefficients and lasing condition, an analytical treatment of waveguiding followed by specifics for the GaN system and numerical simulations for determining the field distribution, loss, and gain cavity modes pertaining to semiconductor lasers are given. An ample fundamental treatment of spontaneous emission, stimulated emission, and absorptions and their interrelations in terms of Einstein's coefficients and occupation probabilities are given. This treatment segues into the extension of the gain discussion to a more realistic semiconductor with a complex valence band such as that of GaN. The results from numerical simulations of gain in GaN quantum wells are discussed, as well as various pathways for lasing such as electron-hole plasma and exciton-based pathways. Localization, which is very pervasive in semiconductors that are yet to be fully perfected, is discussed in the light of laser operation. Turning to experimental measurements, the method for gain measurement, use of various laser properties such as the delay on the onset of lasing with respect to the electrical pulse, dependence of laser threshold on cavity length to extract important parameters such as efficiency are discussed. The aforementioned discussions culminate in the treatment of performance of GaN-based lasers in the violet down to the UV region of the optical spectrum and applications of GaN-based lasers to DVDs along with a discussion of pertinent issues related to the density of storage.

Volume 3, Chapter 3 treats field effect transistor fundamentals that are applicable to any semiconductor materials with points specific to GaN. The discussion primarily focuses on 2DEG channels formed at heterointerfaces and their use for FETs, including polarization effects. A succinct analytical model is provided for calculating the carrier densities at the interfaces for various scenarios and current voltage characteristics of FETs with several examples. The 2-port network analysis, s-parameters, various gain expressions, circuit parameter extraction of equivalent circuit parameters, for both low and high rf power cases, temperature and dispersion

effects are discussed in detail. Experimental performance of GaN-based FETs and amplifiers is then discussed followed by an in-depth analysis of anomalies in the current voltage characteristics owing to bulk and barrier states, including experimental methods and probes used for cataloging these anomalies. This is followed by the employment of field spreading gate plates and associated performance improvements. This segues into the discussion of noise both at the low-frequency end and high-frequency end with sufficient physics and practical approaches employed. The combined treatment of various low-frequency noise contributions as well as those at high frequencies along with their physical origin makes this treatment unique and provides an opportunity for those who are not specialists in noise to actually grasp the fundamentals and implications of low- and high-frequency noise. Discussion of high-power FETs would not be complete without a good discussion of heat dissipation and its physical pathways, which is made available. Unique to GaN is the awareness of the shortfall in the measured electron velocity as compared to the Monte Carlo simulation. Hot phonon effects responsible for this shortfall are uniquely discussed with sufficient theory and experimental data. Power dissipation pathways from hot electrons to hot LO phonons followed by decay to LA phonons and in turn to heat transfer to the bath are treated with sufficient physics. In particular, the dependence of the hot phonon lifetime on the carrier concentration and its implications to carrier velocity is treated. The effect of lattice matched AlInN Barrier layers *vis a vis* AlGaN barrier layers on the hot LO phonon lifetime and carrier velocity is treated. A section devoted to reliability with specifics to GaN based high power HFETs is also provided. Such effects as surfaces, carrier injection by the gate to the surface states and the resultant virtual extended gate, surface passivation, interplay of temperature, strain, and electric field and their combined effect on reliability are treated in detail. Finally, although GaN-based bipolar transistors are not all that attractive at this time, for completeness and the benefit of graduate students and others who are interested in such devices, the theory, mainly analytical, of the operation of heterojunction bipolar transistors is discussed along with available GaN based HBT data.

Volume 3, Chapter 4 discusses optical detectors with special orientation toward UV and solar-blind detectors. Following a discussion of the fundamentals of photoconductive and photovoltaic detectors in terms of their photo response properties, a detailed discussion of the current voltage characteristic of the same, including all the possible current conduction mechanisms, is provided. Because noise and detectors are synonymous with each other, sources of the noise are discussed, followed by a discussion of quantum efficiency in photoconductors and p–n junction detectors. This is then followed by the discussion of vital characteristics such as responsivity and detectivity with an all too important treatment of the cases where the detectivity is limited by thermal noise, shot current noise, generation-recombination current noise, and background radiation limited noise (this is practically nonexistent in the solar-blind region except the man-made noise sources). A unique treatment of particulars associated with the detection in the UV and solar-blind region and requirements that must be satisfied by UV and solar-blind detectors, particularly, for the latter, is then provided. This leads the discussion to various UV detectors

based on the GaN system, including the Si- and SiC-based ones for comparison. Among the nitride-based photodetectors, photoconductive variety as well as the metal-semiconductor, Schottky barrier, and homo- and heterojunction photodetectors are discussed along with their noise performance. Nearly solar-blind and truly solar-blind detectors including their design and performance are then discussed, which paves the way for the discussion of avalanche photodiodes based on GaN. Finally, the UV imagers using photodetectors arrays are treated.

It is fair to state that I owe so much to so many, including my family members, friends, coworkers, colleagues, and those who contributed to the field of semiconductors in general and nitride semiconductors in particular, in my efforts to bring this manuscript to the service of readers. To this end, I thank my wife, Amy, and son, Erol, for at least their understanding why I was not really there for them fully during the preparation of this manuscript, which took longer than most could ever realize. Also, without the support of VCU, with our Dean R. J. Mattauch, Assistant Dean Susan Younce, Department Chair A. Iyer, and my coworkers and students, it would not have been possible to pursue this endeavor. Special recognitions also go to Dr N. Izyumskaya for reading the entire manuscript for consistency in terms of figures, references, and so on, which had to have taken perseverance beyond that many could muster; Dr Ü. Özgür for being the bouncing board and proofing many parts of the book, particularly chapters dealing with optical processes, lasers and magnetism; my colleague P. Jena for reading and contributing to the band structure section; my coworker Professor M. Reshchikov for his contributions to the point defects and doping sections; Professor A. Baski for her expert assistance in obtaining microprobe images; Dr D. Huang for his many contributions to the quantum dots section; Dr Y-T Moon for his assistance in current crowding; C. Liu for her assistance with ferromagnetism; Prof. A. Teke for reading the chapter on detectors; Dr. R. Shimada for her contributions to the surface emitting laser section; Dr. J.-S. Lee for his help in updating the LED chapter; Dr Q. Wang for her help in generating the accurate ball and stick diagrams in Volume 1, Chapter 1; Dr V. Litvinov for calculating the energy levels in quantum wells; students Y. Fu, Fan Qian, X. Ni, and S. Chevtchenko for their contributions to various sections of the book with proofing equations, redoing calculations, and so on; and to J. Leach who took it upon himself to be the local expert in the latest in semiconductor and organic LEDs and helped with the chapter on LEDs and read the chapter on transport as well as proofread some of the other chapters; Ms G. Esposito for reading a large portion of the text for English. Undergraduate students K. Ngandu, D. Lewis, B. D. Edmonds, and M. Mikkelson helped in reading various parts of the manuscript as well as helping with the artwork. Unbeknown to them, many graduate students who took classes from me helped in many immeasurable ways. In terms of the non-VCU colleagues, special thanks go to Professors R. M. Feenstra, A. Matulionis, A. Blumenau, P. Ruterana, G. P. Dimitrakopulos, P. Handel, K. T. Tsen, T. Yao, P. I. Cohen, S. Porowski, B. Monemar, B. Gil, P. Le Febvre, S. Chichibu, F. Tuomisto, C. Van de Walle, M. Schubert, F. Schubert, H. Temkin, S. Nikishin, L. Chernyak, J. Edgar, T. Myers, K. S. A. Butcher, O. Ambacher, A. di Carlo, F. Bernardini, V. Fiorentini, M. Stutzmann, F. Pollak, C. Nguyen, S. Bedair, N. El-Masry, S. Fritsch, M. Grundman, J. Neugebauer, M. S. Shur,

J. Bowers, J. C. Campbell, M. Razhegi, A. Nurmikko, M. A. Khan, J. Speck, S. Denbaars, R. J. Trew, A. Christou, G. Bilbro, H. Ohno, A. Hoffmann, B. Meyer, B. Wessels, N. Grandjean, and D. L. Rode; and Drs Z. Liliental-Weber, P. Klein, S. Binari, D. Koleske, J. Freitas, D, Johnstone, D. C. Look, Z.-Q. Fang, M. MacCartney, I. Grzegory, M. Reine, C. W. Litton, P. J. Schreiber, W. Walukiewicz, M. Manfra, O. Mitrofanov, J. Jasinski, V. Litvinov, Jan-Martin Wagner, K. Ando, H. Saito, C. Bundesmann, D. Florescu, H. O. Everitt, H. M. Ng, I. Vurgaftman, J. R. Meyer, J. D. Albrecht, C. A. Tran, S.-H. Wei, G. Dalpian, N. Onojima, A. Wickenden, B. Daudin, R. Korotkov, P. Parikh, D. Green, A. Hansen, P. Gibart, F. Omnes, M. G. Graford, M. Krames, R. Butte, and M. G. Ganchenkova for either reading sections of the book, providing unpublished data, or providing suggestions. Many more deserve a great deal of gratitude for willingly spending considerable time and effort to provide me with digital copies of figures and high-quality images, but the available space does not allow for individual recognition. They are acknowledged in conjunction with the figures. In a broader sense, on a personal level, it gives me great pleasure to recognize that I benefited greatly from the counsel and support of Professor T. A. Tombrello of Caltech.

I also would like to use this opportunity to recognize a few of the unsung heroes, namely, Dr Paul Maruska and Professor Marc Ilegems who truly started the epitaxy of nitrides with the hydride VPE technique independently, and Dr S. Yoshida and Professor T. Matsuoka for their pioneering work in AlGaN and InGaN, respectively.

Richmond, VA January 2008 *Hadis Morkoç*

Color Tables

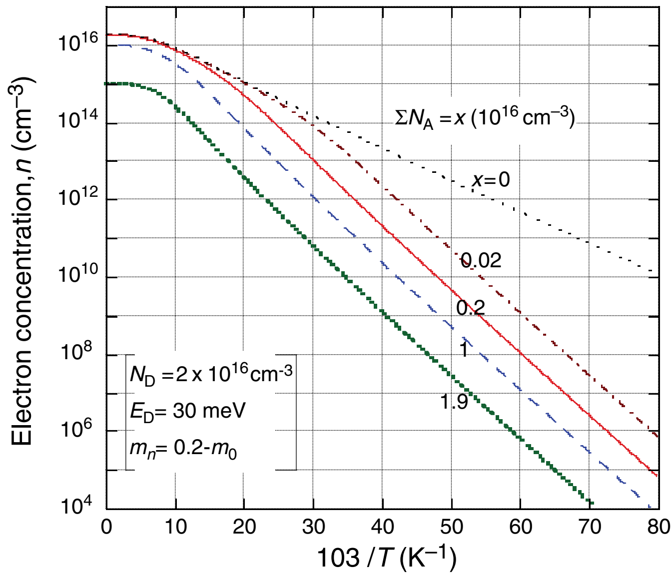


Figure 2.5 Temperature dependence of electron concentration for a donor concentration of $2 \times 10^{16}\,cm^{-3}$, donor binding energy 30 meV, and effective electron mass $0.22m_0$, as a function of temperature for a range of acceptor concentrations from 0 to $1.9 \times 10^{16}\,cm^{-3}$ in GaN. Courtesy of M. Reshchikov.
(This figure also appears on page 139.)

Handbook of Nitride Semiconductors and Devices. Vol. 2. Hadis Morkoç
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40838-2

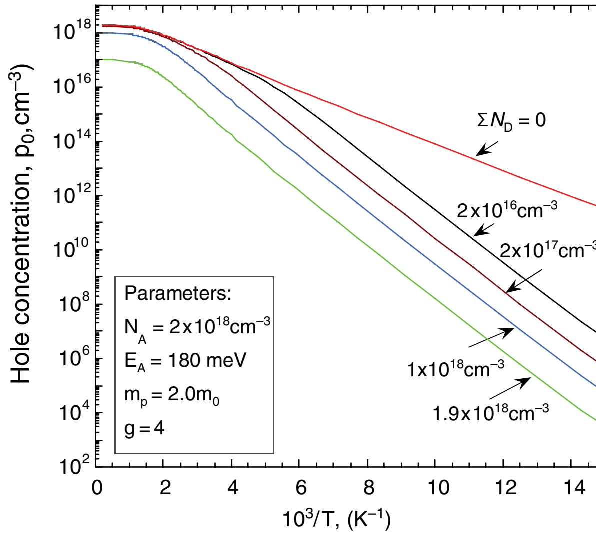


Figure 2.11 Temperature dependence of hole concentration for an acceptor concentration of $2 \times 10^{18}\,\text{cm}^{-3}$, donor binding energy 180 meV, and effective electron mass $2m_0$, as a function of temperature for a range of donor concentrations from 0 to $1.9 \times 10^{18}\,\text{cm}^{-3}$ in GaN. Courtesy of M. Reshchikov and S. S. Chevchenko.

(This figure also appears on page 144.)

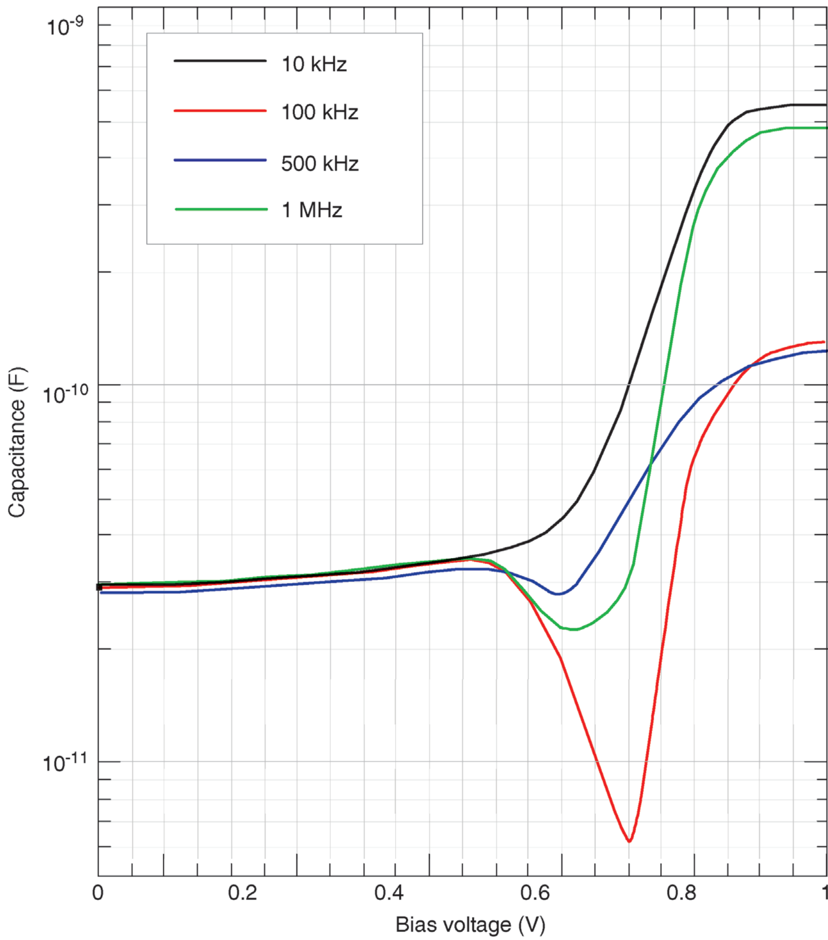


Figure 2.16 Capacitance versus voltage for a GaN Schottky device in the frequency range of 10 kHz–10 MHz.
(This figure also appears on page 156.)

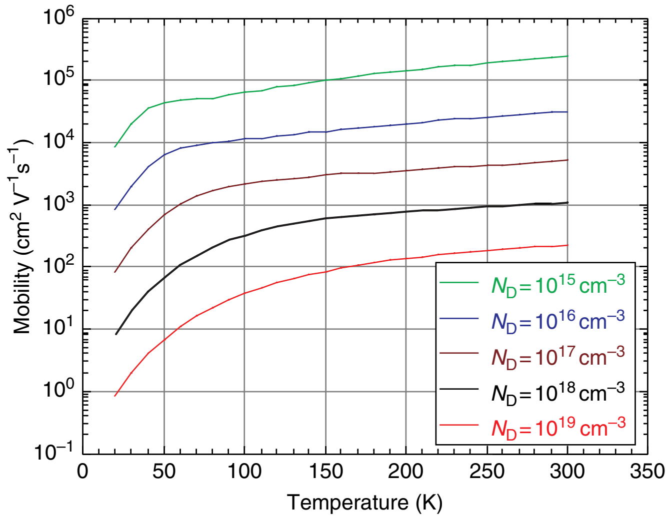


Figure 3.3 Electron mobility calculated for ionized impurity scattering in GaN for a range of ionized impurity concentration from 10^{15} to 10^{19} cm^{-3} with 20% compensation using Equation 3.93.
(This figure also appears on page 189.)

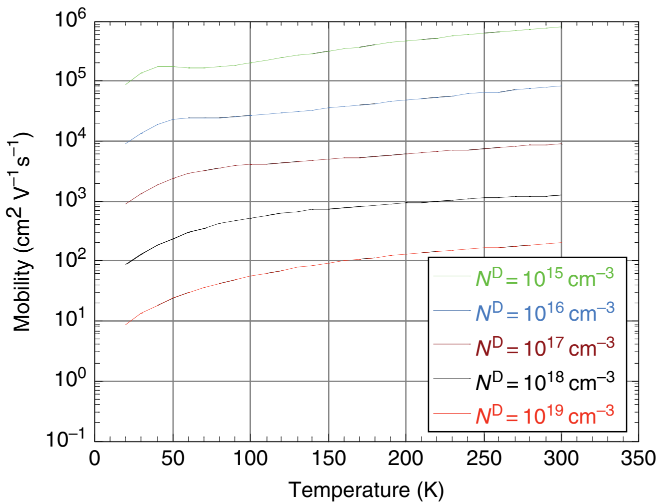


Figure 3.4 Electron mobility calculated for ionized impurity scattering in GaN for a range of ionized impurity concentration from 10^{15} to 10^{19} cm^{-3} with 20% compensation using the phase shift theory.
(This figure also appears on page 189.)

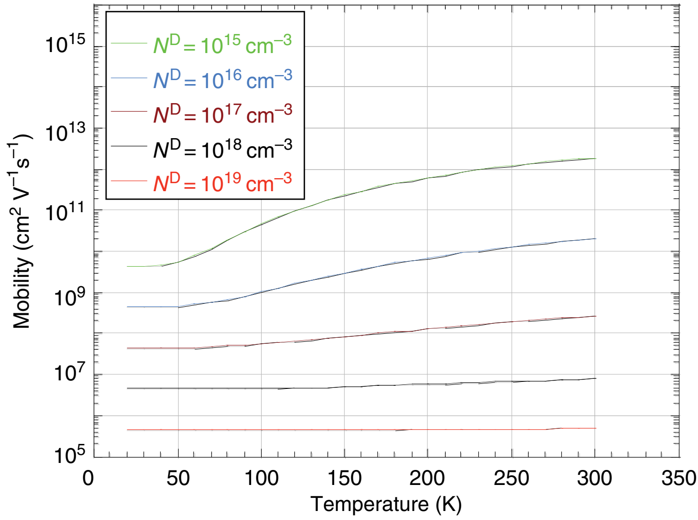


Figure 3.5 Electron mobility in GaN limited by neutral impurity scattering versus temperature for an impurity concentrations in the range of 10^{15}–10^{19} cm^{-3}. Note that this scattering would only be deemed important for impurity concentrations over 10^{21} cm^{-3} at low temperature in high-purity samples.
(This figure also appears on page 192.)

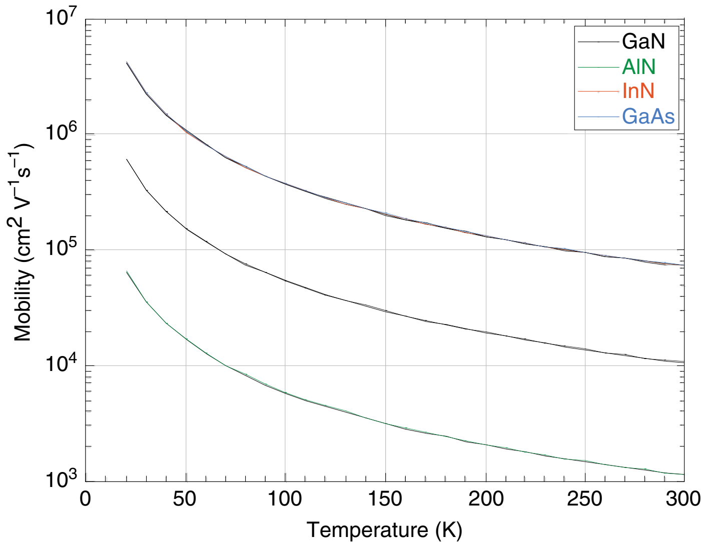


Figure 3.8 Deformation potential (acoustic phonon) scattering limited mobility in InN, GaN, AlN and for comparison in GaAs.
(This figure also appears on page 197.)

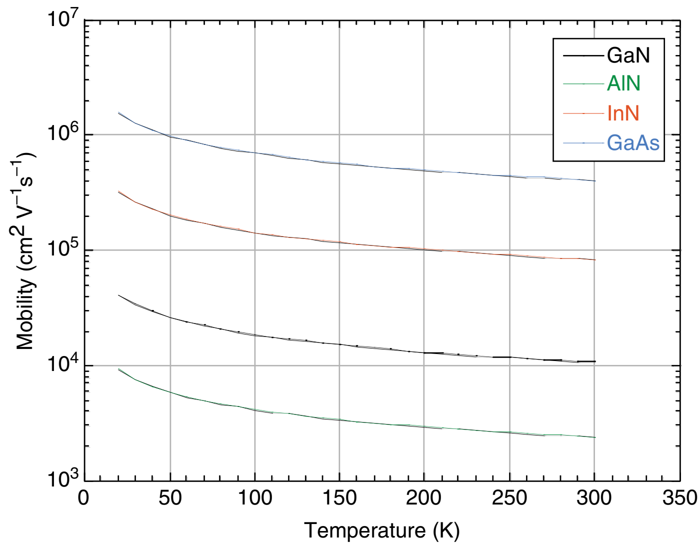


Figure 3.9 Piezoelectric phonon scattering limited mobility in InN, GaN, AlN and for comparison in GaAs.
(This figure also appears on page 200.)

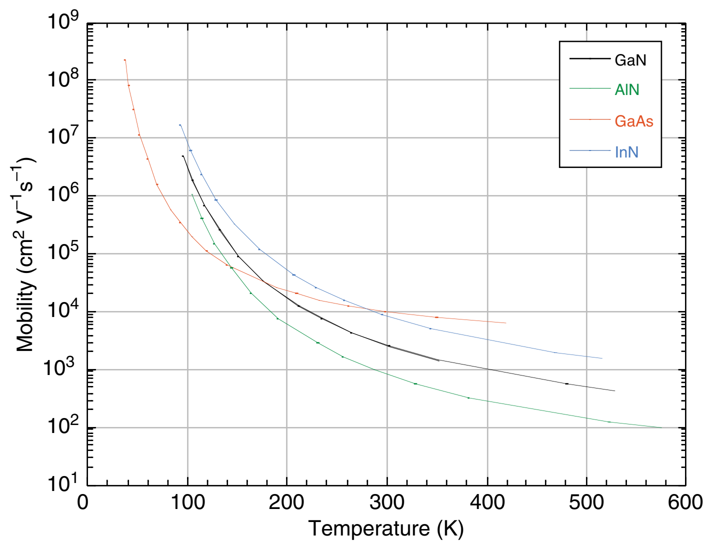


Figure 3.13 Polar optical phonon scattering limited mobility versus temperature for, GaN, AlN, InN and for comparison GaAs using Equation 3.132 with the aid of Equation 3.133.
(This figure also appears on page 210.)

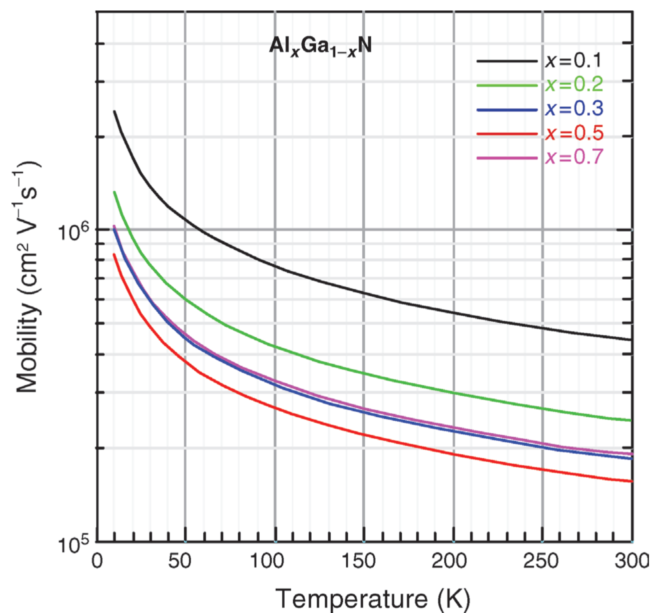


Figure 3.14 The alloy scattering limited mobility in $Al_xGa_{1-x}N$ versus temperature for AlN molar fractions in the range of 0.1–0.7 using Ridley's model described in Equations 3.154 and 3.155. (This figure also appears on page 215.)

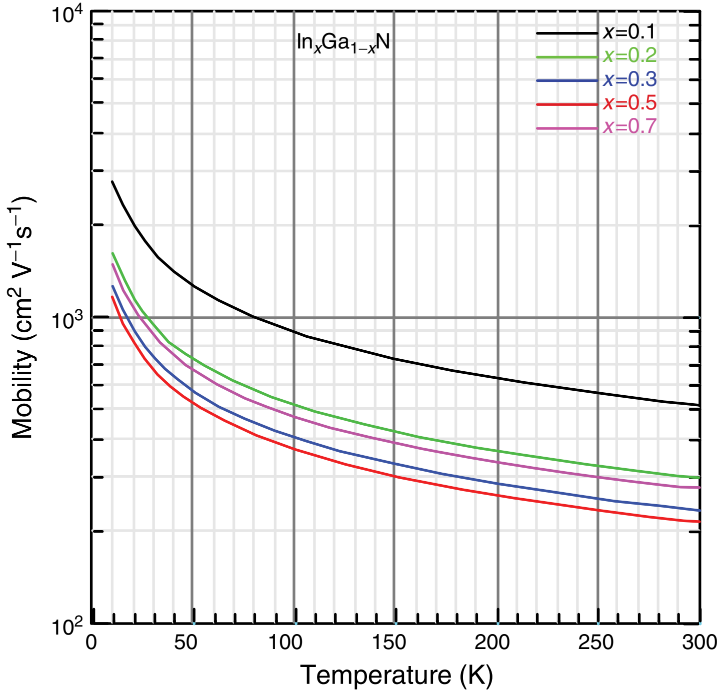


Figure 3.16 The alloy scattering limited mobility in $In_xGa_{1-x}N$ versus temperature for InN molar fractions in the range of 0.1–0.7 using Ridley's model described in Equations 3.154 and 3.155, and an electron effective mass of $0.047m_0$ for InN.
(This figure also appears on page 217.)

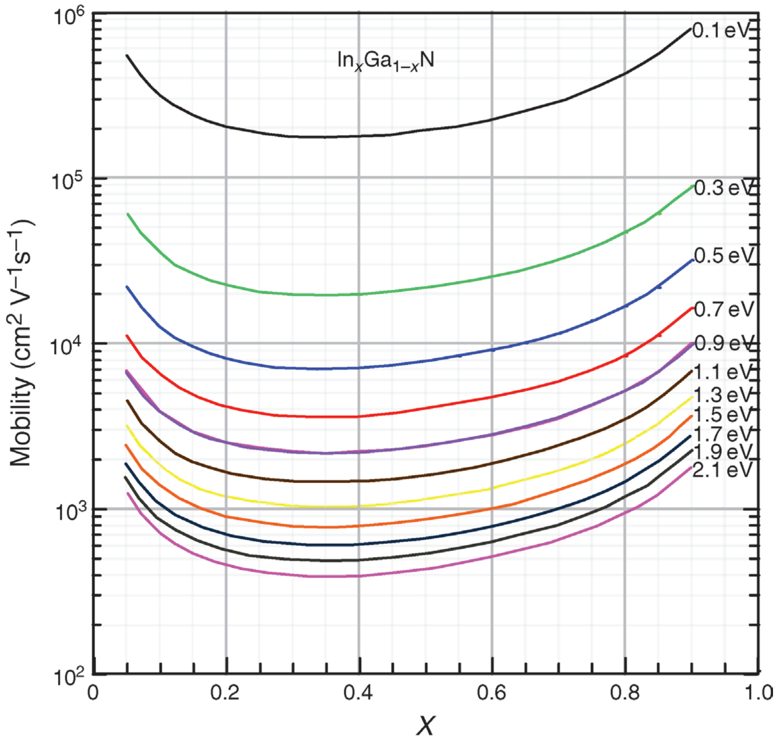


Figure 3.17 The alloy scattering limited mobility for $In_xGa_{1-x}N$ versus the mole fraction with scattering potential ξ_{AB} varied in the range of 0.1–2.1 eV using Ridley's model described in Equations 3.154 and 3.155.
(This figure also appears on page 218.)

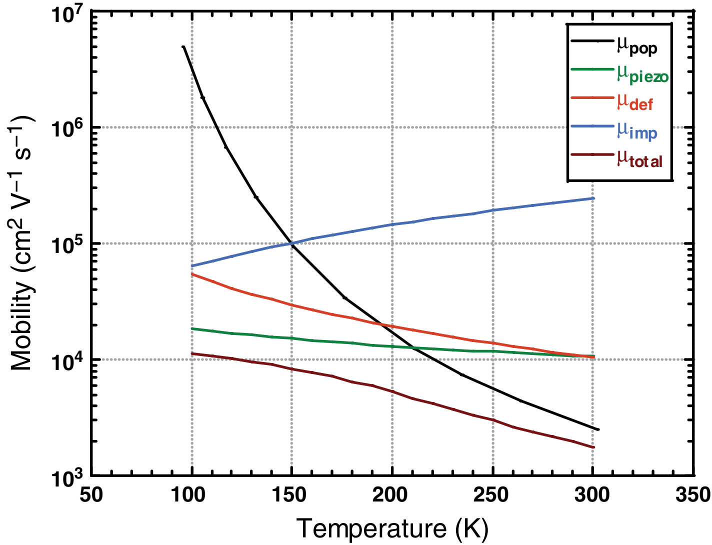


Figure 3.21 Deformation potential scattering, piezoelectric scattering, polar optical phonon scattering, and ionized impurity (10^{15} cm^{-3}) scattering limited mobility in GaN versus temperature along with the cumulative mobility using Matthiessen's rule.
(This figure also appears on page 234.)

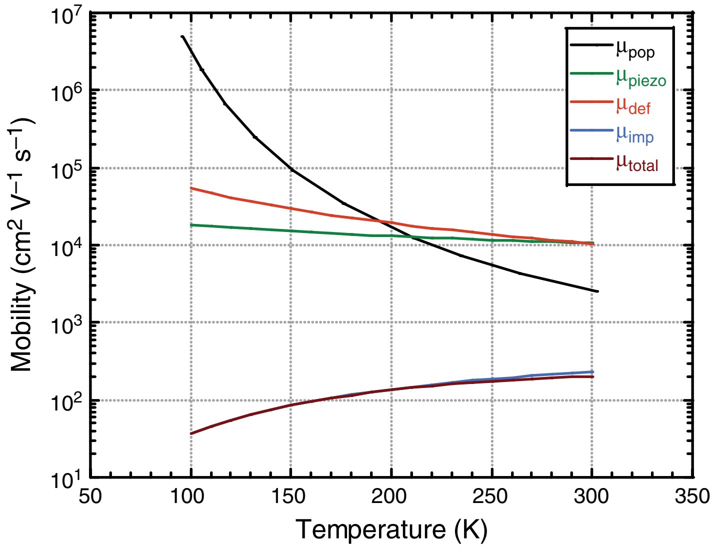


Figure 3.22 Deformation potential scattering, piezoelectric scattering, polar optical phonon scattering, and ionized impurity (10^{19} cm^{-3}) scattering limited mobility in GaN versus temperature along with the cumulative mobility using Matthiessen's rule.
(This figure also appears on page 235.)

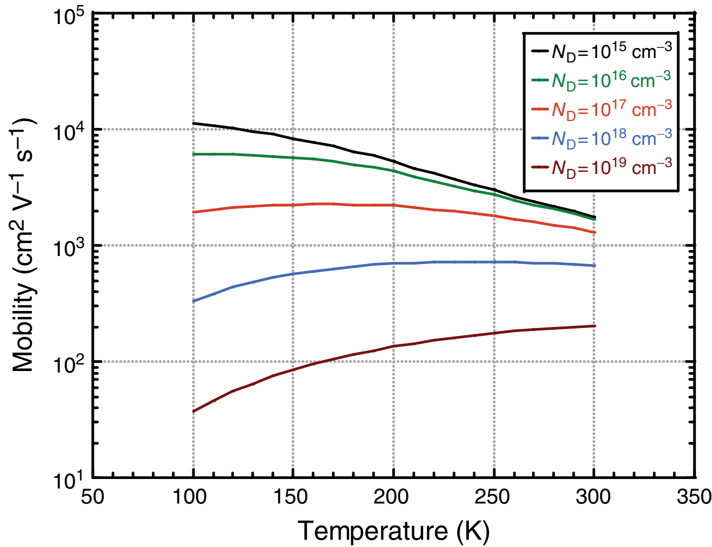


Figure 3.23 Calculated overall electron mobility in GaN versus temperature for ionized donor concentrations of 10^{15}, 10^{16}, 10^{17}, 10^{18}, and 10^{19} cm^{-3} using Matthiessen's rule.
(This figure also appears on page 236.)

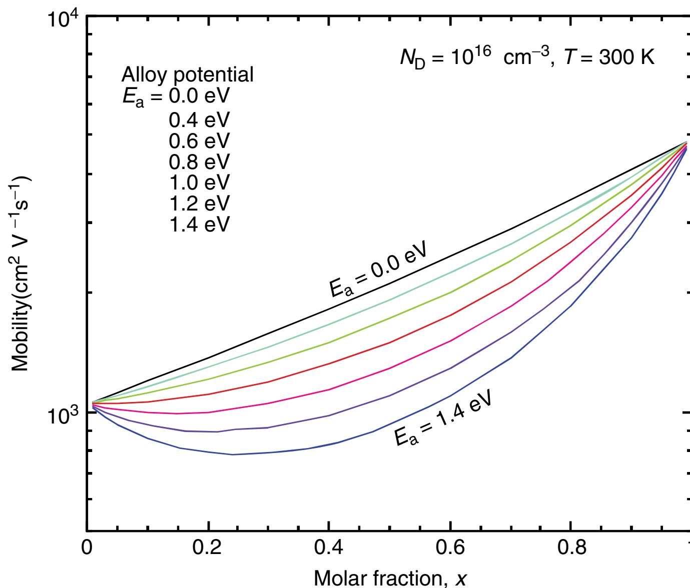


Figure 3.77 Electron mobility for InGaN as a function of the InN molar fraction for an uncompensated donor concentration of 10^{16} cm^{-3} for alloy potentials in the range from 0 to 1.4 V at 300 K.
(This figure also appears on page 347.)

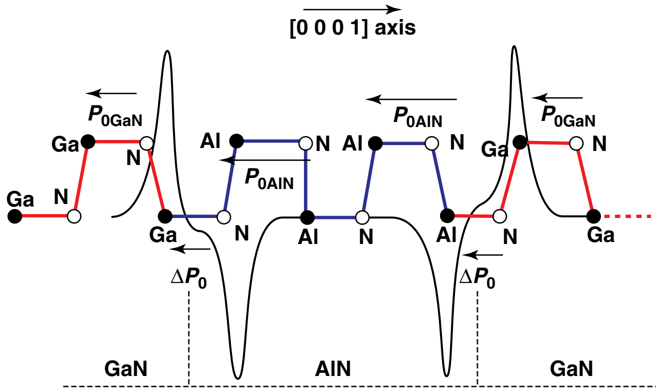


Figure 4.3 The sum of electronic and ionic densities and ensuing electrostatic potential (in Hartree) for an AlN/GaN superlattice-assumed lattice matched to GaN and relaxed. The magnitude of the fields in the bulk regions, meaning not at the interfaces, is 10^7 V cm^{-1}. The horizontal portion of the bonds between the G and N atoms indicate the bonds along the [0 0 0 1] direction. The skewed ones indicate just one of the triple bonds of the tetrahedron [5]. (This figure also appears on page 420.)

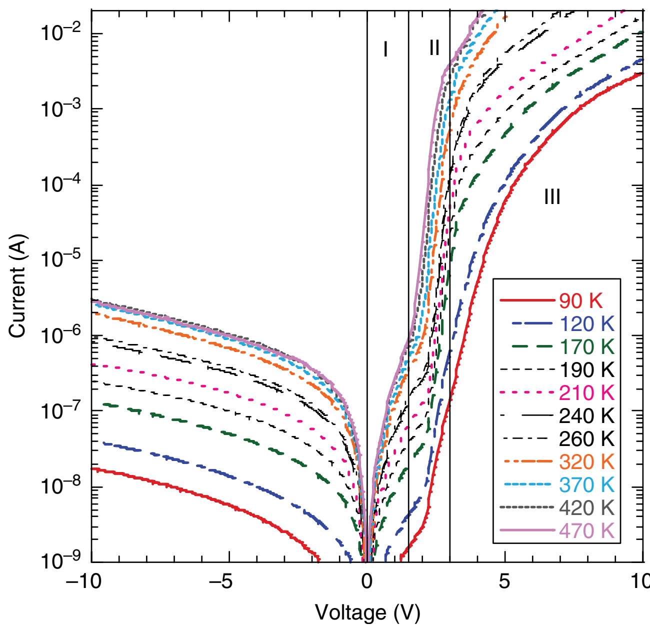


Figure 4.26 Typical temperature-dependent *I–V* curves of 300 × 300 μm p–n junction devices fabricated in GaN grown by MBE for the forward and reverse biases. Three domains namely I, II, and III are delineated and discussed in the text. The temperature legend is such that the highest current in both forward and reverse direction corresponds to the highest temperature, then the next highest, and so on. Courtesy of L. Hirsch [74]. (This figure also appears on page 470.)

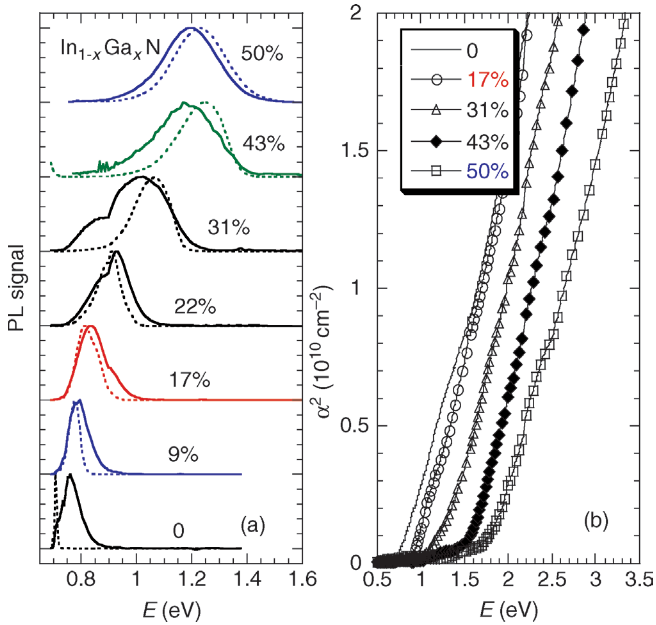


Figure 5.130 (a) PL signal taken at room temperature (solid line) and 11 K (dashed line) for samples with Ga atomic fraction *x* ranging from 0 to 50%. All curves are normalized to equal height and offset vertically for clarity. (b) Room-temperature absorption coefficient squared as a function of photon energy [456].
(This figure also appears on page 662.)

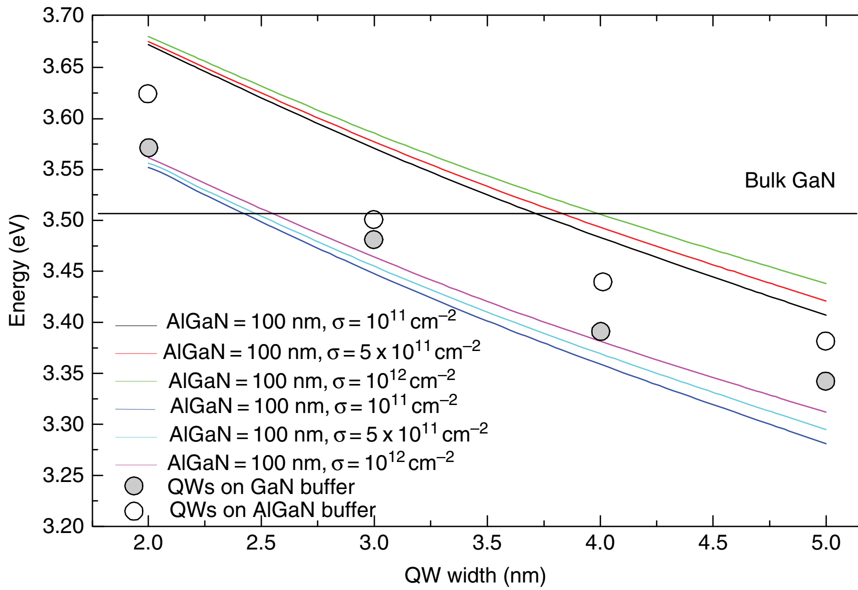


Figure 5.144 The transition energies in GaN/AlGaN quantum wells prepared on GaN and AlGaN buffer layers as a function of well thickness with quantum well charge densities (such as owing to polarization) of 10^{11}, 5×10^{11}, and 10^{12} cm^{-2}. In MQWs on AlGaN buffer layers, composition of the barrier and buffer layers is the same as $Al_{0.15}Ga_{0.85}N$. The solid and open circles indicate results with GaN and $Al_{0.15}Ga_{0.85}N$ buffer layers, respectively. Calculated values for the indicated injection levels (depicted by charge density) are also shown exhibiting very good agreement with experiments.
(This figure also appears on page 684.)

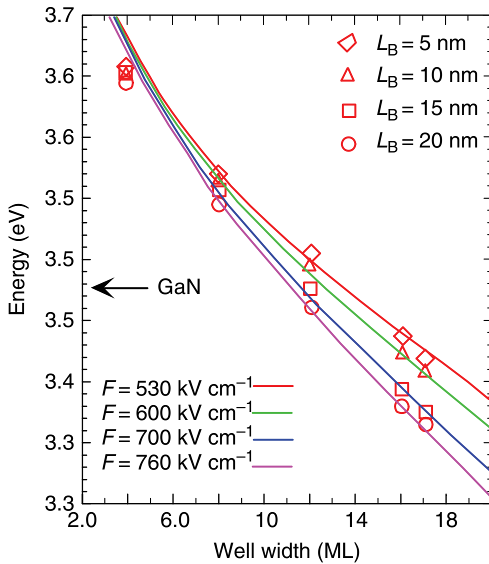


Figure 5.145 The dependence of the transition energies on well width, both experimental and calculated, in $Al_{0.17}Ga_{0.83}N$/GaN QWs as a function of the barrier width and for different electric-field values [492].
(This figure also appears on page 687.)

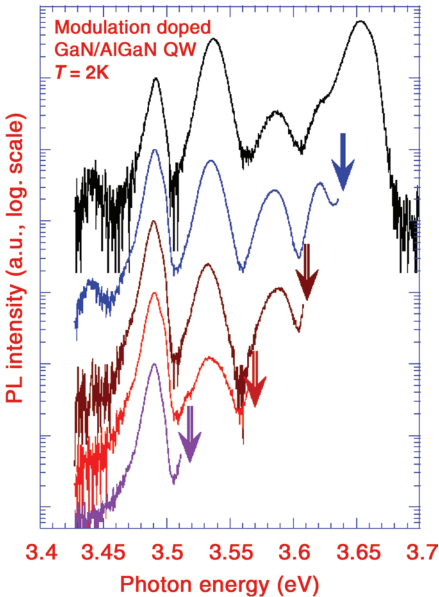


Figure 5.156 Photoluminescence spectra of a modulation-doped GaN/AlGaN QW with different excitation photoenergies. The topmost spectrum was excited with 3.72 eV photons. The three middle peaks originate from emissions in the QW [536].
(This figure also appears on page 703.)

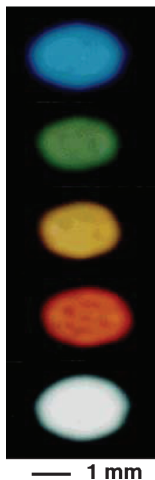


Figure 5.176 Photographs of the light emitted at room temperature from GaN/AlN QDs on Si(1 1 1) substrate excited by a 10 mW unfocused He–Cd laser (∼0.3 W cm^{-2}) [571].
(This figure also appears on page 732.)

1
Metal Contacts to GaN and Processing

Introduction

It is imperative that a semiconductor device be connected to the outside world with no adverse change to its current–voltage characteristics and no additional voltage drop. This can be accomplished only through low-resistance ohmic contacts to the semiconductor. An ideal contact is one where, when combined with the semiconductor, there are no barriers to the carrier flow in either the positive or negative direction. Ideally, this occurs when the semiconductor and the metal work functions are about the same and there are no appreciable interface states, which tend to pin the Fermi level. Because one cannot just dial up ideal work functions for the semiconductor–metal system under consideration, particularly when the work function of the semiconductor varies with doping, it is usually not possible to find just the right combination. In fact, for large-bandgap semiconductors such as GaN, a metal with a large enough work function to form an ohmic contact to p-type GaN does not exist.

As will be clear from the analysis to follow, exacerbating the situation is the large effective mass of the carriers, particularly holes in wide-bandgap semiconductors. Consequently, other options must be explored. Traditionally, these solutions center on increasing the surface doping level and affecting the semiconductor surface through chemical interaction with the metal in a way to render it conducive for current conduction without rectification. Ideally, a metal that is either a donor or an acceptor for n- and p-type contacts, respectively, would be very much in demand. A case in point is Al on Si, which is also a p-type dopant. Another case is the AuGe on GaAs where Ge is an n-type dopant. In addition, ohmic contacts require that contacts are stable both thermally and chemically. The need for stability cannot be overstated in devices intended for high-power and high-temperature operation either by design or necessity when the junction temperatures could be very high. For example, in LEDs, the power loss at the contacts reduces the wall-plug efficiency and increases the junction temperature. This potentially degrades the operating lifetime. In lasers, requiring high current levels to operate particularly in the early stages of development, ohmic contacts may make the difference between a successful attempt and a failed one.

The inability to count on barrier-free contacts for p-type GaN warrants an understanding of the current conduction in metal–semiconductor systems germane

Handbook of Nitride Semiconductors and Devices. Vol. 2. Hadis Morkoç
Copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40838-2

to wide-bandgap semiconductors. Ironically, in the event that the contact to p-type GaN is not ohmic, a forward-biased p–n junction would inevitably transform the metal to a p-GaN contact to a reverse-biased Schottky barrier, making a difficult problem go from bad to worse. In the absence of defects and high surface doping, only those carriers that have sufficient kinetic energy to surmount the barrier, which is described by thermionic emission (TE) and field emission (FE), would contribute to the current flow and power dissipation. In the presence of defects, and in the event of high surface doping resulting in thin depletion layers, defect-assisted tunneling, field-aided tunneling, and direct tunneling must also be considered.

1.1 A Primer for Semiconductor–Metal Contacts

When a metal and a semiconductor with no surface states are brought in contact and equilibrium is maintained, their Fermi levels will align. If the Fermi levels of the metal and semiconductor were the same as before the contact, then there would be no change in the band structure after contact. Because the Fermi level in the semiconductor, and thus the work function, depends on carrier concentration, matching the work functions is nearly impossible, even though attempting to do so would be better than not.

Let us consider the case of an n-type semiconductor and a metal with a work function that is larger than that of the semiconductor. The alignment of the Fermi levels after contact, brought about by the charge motion from the higher toward the lower energy side, creates a depletion region in the semiconductor and a barrier at the interface. The barrier height ϕ_B (before the image force lowering discussed below) is simply the difference between the metal work function ϕ_m and the electron affinity in the n-type semiconductor $(\phi_m - \chi)$, as shown in Figure 1.1. In this ideal picture, the band bending in the semiconductor is simply the difference between the metal work function ϕ_m and semiconductor work function ϕ_s $(\phi_m - \phi_s)$.

The above simple picture illustrates the image force lowering. This lowering comes about from a negative charge at a distance x from the surface of, say a metal for convenience, inducing a positive charge of equal value at a distance $-x$ from the surface. The confining barrier can be lowered by application of an electric field normal to the surface as shown in Figure 1.2. This is called the image force lowering.

The attractive force between an electron and a positive image force in the metal separated by a distance $2x$ (x is the distance between the electron under consideration and the metal surface and $-x$ is the distance from the metal surface to the metal of the image charge) is given by

$$F = \frac{-q^2}{4\pi\varepsilon_0 (2x)^2}. \tag{1.1}$$

The terms have their usual meanings. The work done by the same force is the integral of the force over distance, x, from 0 to x:

$$E(x) = -\frac{q^2}{16\pi\varepsilon_0 x}, \tag{1.2}$$

which is represented by the uniform dashed line in Figure 1.2.

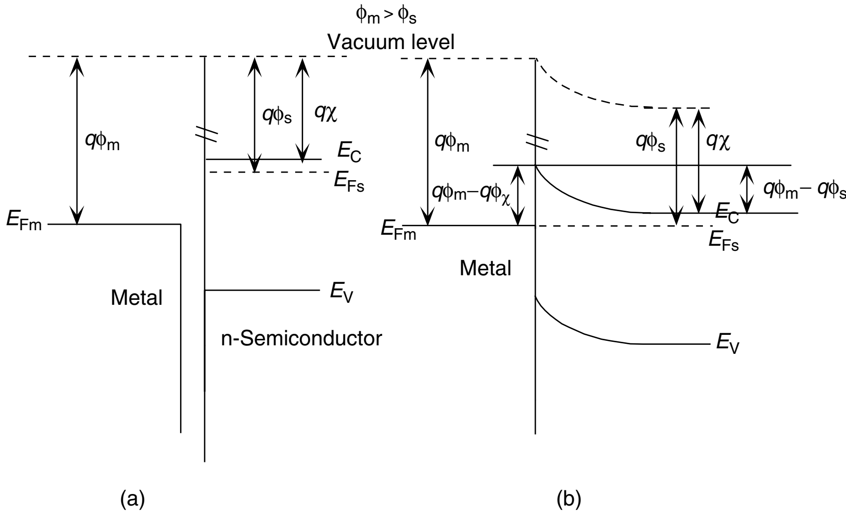


Figure 1.1 A metal n-type semiconductor pair before (a) and after (b) contact with no surface/interface states. The metal work function is greater than that for the semiconductor ($\phi_m > \phi_s$).

With an applied electric field, E, applied normal to the metal surface, the total potential energy is the sum of that given by Equation 1.2 and qEx and given as

$$\mathrm{PE}(x) = -\frac{q^2}{16\pi\varepsilon_0 x} - qEx. \tag{1.3}$$

The potential energy is shown in Figure 1.2 as solid lines for three different values of the applied electric field with solid lines. Note that a built-in electric field normal to the surface at the interface already exists after contact. Therefore, the equations above are valid also for the case where the vacuum is replaced by the semiconductor.

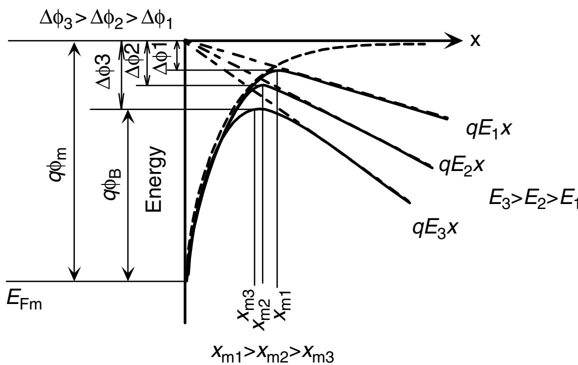


Figure 1.2 Image force lowering for an electron at a distance x away from a metal surface for three different electric fields, $E_1 < E_2 < E_3$. The terms x_i ($i = 1, 2, 3$) indicate the maximum of the total energy. $\Delta\phi_i$ ($i = 1, 2, 3$) are the image force lowering of the barrier.

Application of an additional electric field (bias) will contribute to the total electric field appearing in Equation 1.3. The maximum of the potential energy occurs at x_m, (minimum in the case of p-type semiconductor), which is where the derivative of Equation 1.3 goes to zero. Doing so leads to

$$x_m = \sqrt{\frac{q}{16\pi\varepsilon_s E}}. \tag{1.4}$$

Substituting Equation 1.4 into Equation 1.3 leads to the image force lowering $\Delta\phi$ as

$$\Delta\phi = \sqrt{\frac{qE}{4\pi\varepsilon_s}}. \tag{1.5}$$

Additional treatment can be found in many texts [1,2]. Clearly, the image force lowering increases through a square root dependence with the electric field. This means that the effective barrier for electron escaping the metal is given by the difference between the metal work function and the image force lowering, expressed as

$$\phi_B = \phi_m - \Delta\phi. \tag{1.6}$$

The equilibrium case of Figure 1.6a, where no external bias is applied to the metal–semiconductor rectifying contact, forward bias (negative voltage applied to the n-type metal with respect to the metal) lowers the barrier by the applied bias, as shown in Figure 1.3a, and a reverse bias (positive voltage applied to the n-type semiconductor with respect to the metal) increases the barrier, as shown in Figure 1.3b. The barrier to electron flow from the metal to the semiconductor remains almost unchanged except through a change in $\Delta\phi$. This image force lowering term increases with increasing electric field because of a reverse bias and decreases with forward bias. In addition, the

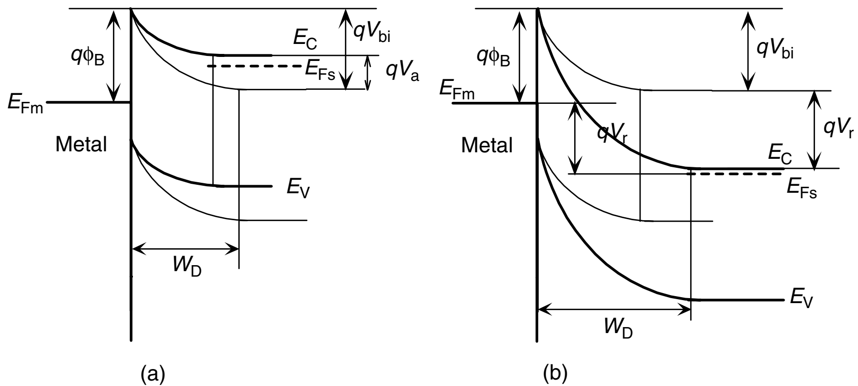


Figure 1.3 A metal n-type semiconductor system where the metal work function is greater than the semiconductor work function in forward-bias (a) and reverse-bias (b) cases. The terms V_a and V_r represent the forward- and reverse-bias voltages, respectively. The extension of depletion regions (W_D) is also shown so is the equilibrium band diagram of the semiconductor in lighter pen. The image force lowering is not shown for simplicity.

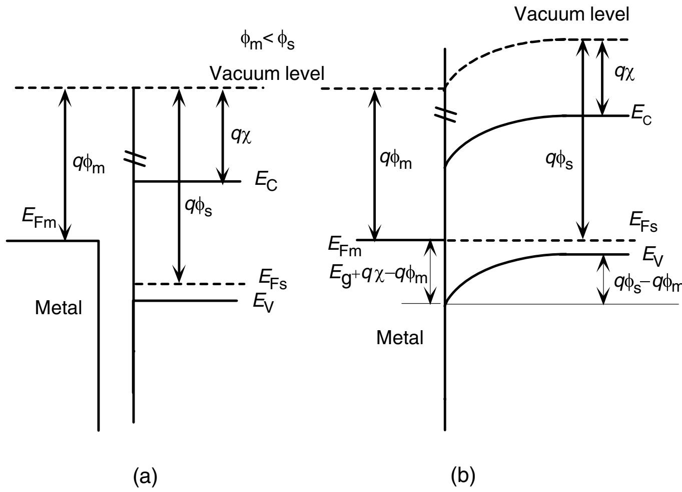


Figure 1.4 A metal p-type semiconductor pair before (a) and after (b) contact with no surface/interface states. The metal work function is smaller than that for the semiconductor ($\phi_m < \phi_s$).

position of the peak in the barrier, x_{m}, moves closer to the metal– semiconductor interface with increasing electric field owing to the reverse bias.

The band diagram for a rectifying metal p-type semiconductor system before (a) and after (b) contact (in equilibrium) is shown in Figure 1.4. The same with forward (positive voltage applied to the semiconductor with respect to the metal) is shown in Figure 1.5a and in reverse bias (negative voltage applied to the semiconductor with respect to the metal) is shown in Figure 1.5b.

Returning to an ideal but rectifying metal n-type semiconductor contact with the image force lowering and conduction band only is shown in Figure 1.6a–c for equilibrium forward- and reverse-bias cases. Note that the image force lowering term $\Delta\phi$ is a function of bias increasing with reverse bias and decreasing with forward bias as determined by Equation 1.5. The picture for a metal p-type semiconductor system is similar to that for the n-type case of Figure 1.6 and is illustrated in Figure 1.7 for the equilibrium and reverse-bias cases. The reverse bias is chosen as it depicts precisely what could take place in a forward-biased contact p–n junction, where the contact to the p-type semiconductor is nonohmic (and reverse biased for the polarity making the p–n junction forward biased), as is often the case for wide-bandgap semiconductors unless the hole concentration is high. Dependencies of the barrier lowering and the position of the potential maximum on the applied bias presented for the n-type case apply here also. As indicated in Figure 1.7, the current conduction can be owing to defect-assisted tunneling, thermionic field emission (TFE), and thermionic emission. An ideal direct tunneling is unlikely considering the effective mass of the holes, the low hole concentration, and a barrier height that is most likely high. In fact, the current conduction in GaN-based p–n junctions is generally ill understood and may involve band tail states, a topic of discussion in Chapters 4 and 5.

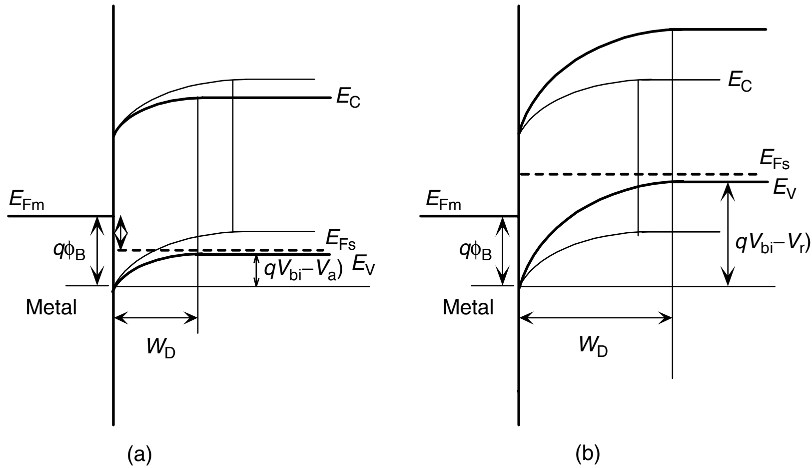


Figure 1.5 A metal p-type semiconductor system where the metal work function is greater than the semiconductor work function in forward-bias (a) and reverse-bias (b) cases. The terms V_a and V_r represent the forward- and reverse-bias voltages, respectively. The extension of depletion regions (W_D) is also shown so is the equilibrium band diagram of the semiconductor in lighter pen. The image force lowering is not shown for simplicity.

The case where the fortuitous matching of the metal–semiconductor pair occurs is depicted in Figure 1.8 with an ohmic contact behavior that results automatically. The condition here is that the metal work function is equal to or slightly smaller than that of the semiconductor. Unlike the case of Figure 1.6, in which the metal work function is smaller than that of the semiconductor, charge accumulation rather than depletion occurs with a negligible voltage drop. The same is displayed in Figure 1.9 for a p-type semiconductor where the metal work function needs to be equal to or larger than that for the n-type semiconductor. For p-type GaN, this would mean a metal with a work function of about 8 eV; however, this does not exist. Together with a large hole mass and the difficulty of obtaining high hole concentrations, this paints a very dismal picture regarding ohmic contacts to p-type GaN.

1.2
Current Flow in Metal–Semiconductor Junctions

When a metal is brought in contact with a semiconductor, there arises a potential barrier unless the work functions of the metal and the semiconductor match, which is highly unlikely. If a bias is applied, the current flow takes place when the carriers in the metal or in the semiconductor gain sufficient energy, by thermal means or by field, to overcome the barrier. However, when the barrier is sufficiently thin, they can also go through the barrier either by direct tunneling if the barrier thickness is comparable with the tunneling distance or by gaining sufficient energy with respect to

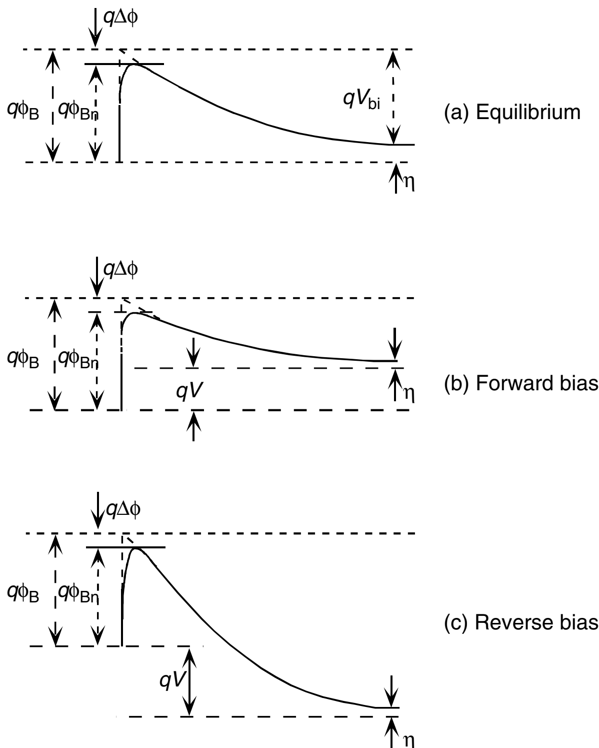


Figure 1.6 (a–c) Ideal metal n-type semiconductor contacts under equilibrium, forward, and reverse bias. Also shown is the image force lowering of the barrier.

the Fermi level combined with tunneling at some point in the barrier. The current conduction process over or through a barrier created by a metal–semiconductor contact is schematically shown in Figure 1.10. These processes can be circumvented by defects. In cases where defects are not involved, there are three mechanisms [1,2] that govern the current flow in a metal–semiconductor system, which are discussed below.

1. *Thermionic emission.* For lightly or moderately doped semiconductors, $N_D < \approx 10^{17}\,\mathrm{cm}^{-3}$, the depletion region is relatively wide. It is, therefore, nearly impossible for electrons to tunnel through the barrier unless aided by defects, which are considered not to exist in this ideal picture. In a forward-biased junction, however, the electrons can surmount the top of the barrier, which is lowered with respect to the Fermi level in the semiconductor by an amount equal to the applied bias. This is called the thermionic emission as shown in Figure 1.10a and has been treated in many papers and early texts such as that by Henish [3]. In reverse bias, the barrier for electrons from the semiconductor to the metal is made even larger and the electron flow from the semiconductor to the metal in this ideal picture is cut off. On the metal side, if the electrons in the metal gain sufficient energy by the applied bias, they too can overcome the barrier, a dominant mechanism for the reserve

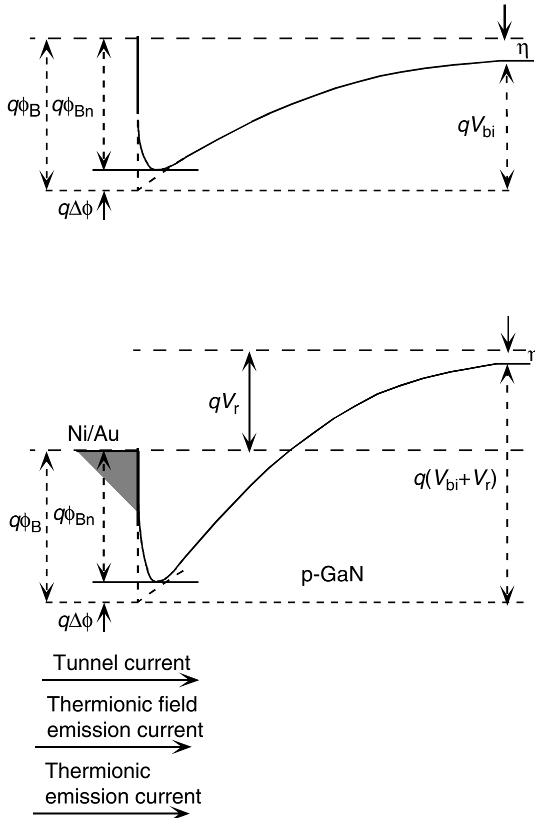


Figure 1.7 Ideal metal p-type semiconductor contacts under equilibrium and reverse bias. Also shown is the image force lowering of the barrier and the various current conduction mechanisms that may come into play.

bias current in an ideal situation. Naturally, an ohmic behavior is not observed. The electron flow from the metal to the semiconductor and vice versa must balance for zero bias under steady state conditions, which ensures zero net current. The thermionic process requires electrons to gain considerable energy for current flow from which it gets its name.

2. *Thermionic field emission*. For intermediately doped semiconductors, $\approx 10^{17} < N_D \approx 10^{18}$ (cm^{-3}), the depletion region is not sufficiently thin to allow direct tunneling of carriers that are more or less in equilibrium. This process requires some energy gain from the bias sufficient to raise the electron energy to a value E_m where the barrier is sufficiently thin for tunneling, as shown in Figure 1.10b. Discussed first by Dolan and Dyke [4] in conjunction with field emission from a metal tip, this process is incorporates the elements of thermionic emission, in the sense that electrons must be moderately hot or warm, and tunneling that requires penetration through a sufficiently thin barrier.

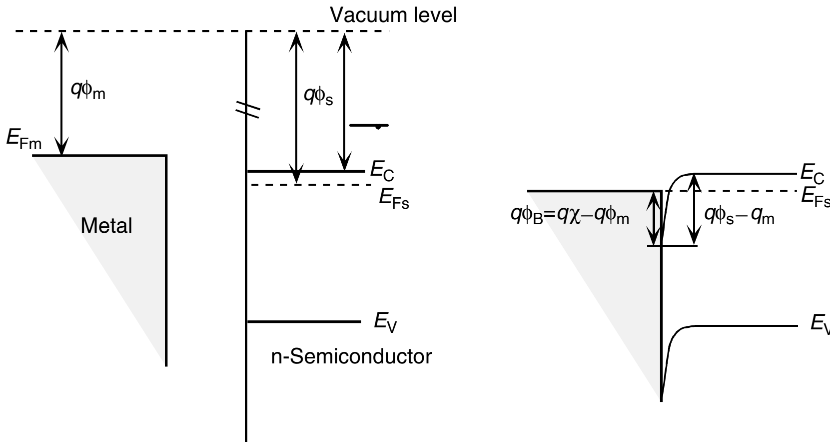


Figure 1.8 Fortuitous (imaginary not real) matching of the metal–n-type semiconductor pair with automatic ohmic contact behavior. The results are brought about by assuming that the metal work function is equal to or smaller than that for the semiconductor (left: before contact; right: after contact).

3. *Field emission*. In heavily doped semiconductors, $N_D > \approx 10^{18}\,cm^{-3}$, the depletion region is narrow even for cold and cool electrons at the bottom of the conduction band or at the Fermi level, the latter is for degenerate semiconductors, and direct electron tunneling from the semiconductor to the metal is allowed as shown in

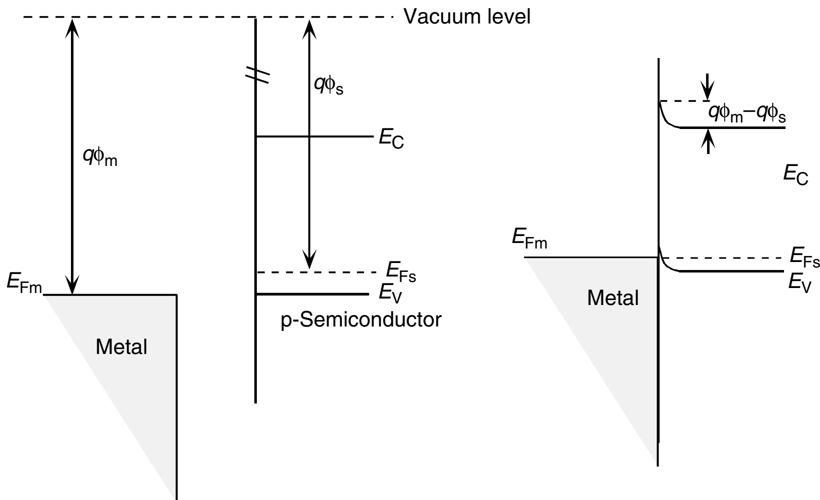


Figure 1.9 Fortuitous (imaginary not real) matching of the metal p-type semiconductor pair with automatic ohmic contact behavior brought about by the assumed metal work function being equal to or greater than that for the semiconductor (left: before contact; right: after contact).

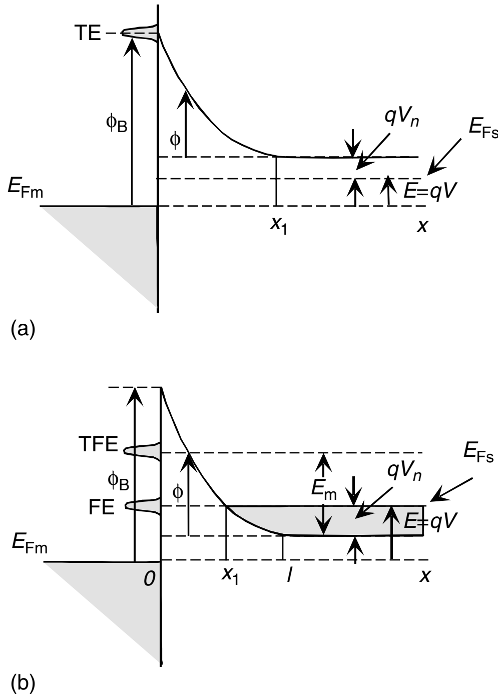


Figure 1.10 Potential energy diagram and current flow mechanisms for a forward-biased Schottky barrier: (a) for the thermionic emission process, which is more likely when the doping level in the semiconductor is relatively low and the Fermi level is below the conduction band, and (b) for thermionic field emission and direct tunneling, also referred to as field emission, which is more likely when the doping level in or on the semiconductor surface is sufficiently high to the extent that the Fermi level may even be in the conduction band as depicted.

Figure 1.10b. In the absence of a good match between the metal and the semiconductor work functions, which is generally the case, this is the best approach to pursue ohmic contacts provided, of course, that very large doping concentrations can be attained. A good discussion of these three processes, TE, TFE, and FE, can be found in Stratton [5].

1.2.1 The Regime Dominated by Thermionic Emission

The traditional current–voltage expression representing thermionic emission is given by

$$J_{\mathrm{te}} = J_{\mathrm{te}\,0}\left[\exp\left(\frac{qV}{kT}\right) - 1\right], \tag{1.7}$$

with

$$J_{te\,0} = A^* T^2 \exp\left[\frac{-q(\phi_B - \Delta\phi)}{kT}\right], \tag{1.8}$$

where $J_{te\,0}$ is the saturation value of the current density J_{te}, A^* is the effective Richardson constant, ϕ_B is the barrier height, and $\Delta\phi$ is the image force barrier lowering. Equation 1.7 is based on the condition that the series resistance of the circuit is negligibly small. It should be pointed out that the saturation current density is typically designated by J_s in general. We do the same throughout most of this chapter. However, in this subsection, to make the point that we are discussing thermionic emission, a more descriptive nomenclature, $J_{te\,0}$, is used.

As the kT term in the exponent indicates, the slope of $J_{te\,0}/T^2$ would vary with temperature with a slope of kT in a semilogarithmic plot. The Richardson constant for free space is given by

$$A^*_{free} = \frac{4\pi q k^2 m_0}{h^3}, \tag{1.9}$$

which equals 120 ($A\,cm^{-2}\,K^{-2}$). The effective Richardson constant is $A^* = A^*_{free}\left(m^*_e/m_0\right)$ for n-type and $A^* = A^*_{free}\left(m^*_{hh}/m_0\right)$ for p-type semiconductors under the assumption of single-valley conduction bands such as n-type GaN and single and spherical valence band conduction. When both heavy- and light-hole bands are occupied, the effective Richardson constant is given by $A^* = A^*_{free}\left(m^*_{hh} + m^*_{lh}/m_0\right)$. In cubic compound semiconductors, the valence band is degenerate and thus the last expression for the Richardson constant should be used.

Equation 1.7 represents the carrier flux from the semiconductor to the metal, with the barrier depending on voltage, $\phi_B - V$; and if from the metal to the semiconductor with the barrier fixed at ϕ_B, there exists a parasitic resistance in the circuit such as semiconductor resistance, the thermionic emission current expression is modified as

$$J_{te} = J_{te\,0}\left[e^{q(V - IR_s)/kT} - 1\right]. \tag{1.10}$$

Here the current, I, is determined by the product of the current density J and the area of the structure.

Because both A^* and $\Delta\phi$ are voltage dependent, it is customary to represent the current–voltage characteristics for applied voltages $>3kT/q$ for simplicity as $J \sim \exp(qV/nkT)$ with n denoting the ideality factor.

In the reverse direction, the barrier lowering becomes more important. In a such a case (using J_s instead of $J_{te\,0}$ for saturation current),

$$J_R \approx J_S = A^* T^2 \exp\left(\frac{-q\phi_B}{kT}\right)\exp\left(\frac{q\Delta\phi}{kT}\right), \tag{1.11}$$

where the image force barrier lowering

$$\Delta\phi = \sqrt{\frac{qE}{4\pi\varepsilon_s}}, \tag{1.12}$$

and the electric field at the metal–semiconductor interface is given by

$$E = \sqrt{\frac{2qN_D}{\varepsilon_s}\left(-V + V_{bi} - \frac{kT}{q}\right)}. \tag{1.13}$$

Neglecting the image force lowering, the barrier height, and the effective Richardson constant can be experimentally determined by plotting $\ln(J_R/T^2)$ versus $1000/T$ (Richardson plot). Actually, the result would be that of the effective barrier height including the image force lowering. If the interface electric field can be determined, the image force lowering can be calculated. Performing the measurements for a range of reverse-bias conditions, which would help determine the saturation current for that range of biases, would allow the determination of the image force lowering component. The assumption here is that components of the current other than the thermionic emission are nonexistent or can be separated out. Confidence can be gained if the image force lowering so determined is linearly dependent on the square root of the interface electric field. From an experimental point of view, generation–recombination current would also increase, unless negligible, which would exacerbate the determination of barrier lowering by image force.

The generation–recombination current, discussed in more detail in Chapter 4, is given by

$$J_{gr} = \frac{qn_i W}{\tau}\exp\left(\frac{qV}{2kT}\right), \tag{1.14}$$

where n_i is the intrinsic concentration, τ is the effective carrier lifetime, and W, the depletion depth, is given by

$$W = \sqrt{\frac{2\varepsilon_s(V_{bi} - V)}{qN_D}}. \tag{1.15}$$

Because the intrinsic carrier concentration is nearly nil at room temperature for GaN, the generation–recombination current component should be negligibly small. Contribution by generation–recombination current to the overall current is depicted in Figures 4.21 and 4.22, and even for very small effective carrier lifetimes, the generation–recombination current is small. Even at elevated temperatures, this remains to be the case because intrinsic concentration at those temperatures remains negligibly small.

1.2.2
Thermionic Field Emission Regime

TFE in classical treatments is assumed to be associated with the intermediate temperature range and where the electrons tunnel from the semiconductor to the

metal at an energy E_m above the conduction band edge, as shown in Figure 1.10b. The component of the current for the TFE process from the semiconductor to the metal for this form of current transport has been expressed by Stratton [6], and Padovani and Stratton [7] as

$$J_{tfe} = \frac{A^* T^2}{2\pi kT}\left(\frac{\pi}{f_m}\right)^{1/2} \exp\left(\frac{qV_n}{kT} - b_m - c_m E_m\right)\left[1 + \operatorname{erf}(E_m f_m^{1/2})\right]. \qquad (1.16)$$

The constants b_m, c_m, and f_m are the Taylor expansion coefficients for the exponent of the transparency of the barrier around an energy E_m. The energy E_m is chosen to satisfy

$$c_m kT = 1. \qquad (1.17)$$

(This would make $c_m E_m$ term in Equation 1.16 E_m/kT.)

If we take the extension of the Fermi level into the conduction band as qV_n as positive, b_m, c_m, and f_m constants are defined as

$$b_m = \frac{[q(\phi_B - V + V_n)^{1/2}(\phi_B - V + V_n - E_m/q)^{1/2} - E_{00}E_m/kT]}{E_{00}}. \qquad (1.18)$$

and

$$c_m = \frac{1}{E_{00}} \log\left[\frac{q^{1/2}(\phi_B - V + V_n)^{1/2} + (q\phi_B - qV + qV_n - E_m)^{1/2}}{E_m^{1/2}}\right], \qquad (1.19)$$

and

$$f_m = \frac{\cosh^2(E_{00}/kT)}{4E_{00}q(\phi_B - V + V_n)}, \qquad (1.20)$$

where

$$E_{00} = \frac{q\hbar}{2}\sqrt{\frac{N_D}{\varepsilon_s m^*}}. \qquad (1.21)$$

The energy E_m at which the electron emission takes place can be found using Equations 1.17 and 1.19 and is given by

$$E_m = \frac{q(\phi_B - V + V_n)}{\cosh^2(E_{00}/kT)}. \qquad (1.22)$$

Padovani and Stratton have shown that the energy distribution of the emitted electron is actually a Gaussian distribution having a half width of

$$\Delta = \sqrt{\frac{\log 2}{f_m}}. \qquad (1.23)$$

An inspection of Equation 1.16 together with Equations 1.18–1.20 leads to the recognition that the current–voltage characteristic is dominated by the exponential factor and its evaluation reveals that

$$b_m + \frac{E_m}{kT} = \frac{q(\phi_B - V + V_n)}{E_0}, \tag{1.24}$$

where

$$E_0 = E_{00} \coth\left(\frac{E_{00}}{kT}\right). \tag{1.25}$$

Neglecting the error function term in Equation 1.16 and employing the terms defined above the forward current density owing to TFE can be expressed as

$$J_{\mathrm{tfeF}} = J_{\mathrm{SF}} \exp\left(\frac{qV}{E_0}\right), \tag{1.26}$$

where J_{SF} is the saturation value of the current J_{tfeF} and expressed by

$$J_{\mathrm{SF}} = \frac{A^* T^2 \sqrt{\pi q E_{00}(\phi_B - V + V_n)}}{kT \cosh(E_{00}/kT)} \exp\left(\frac{qV_n}{kT} - \frac{q(\phi_B + V_n)}{E_0}\right). \tag{1.27}$$

Considering the electron emission from the metal to the semiconductor at energy E_m, the total current in the forward direction and neglecting the error function term in Equation 1.16, the forward current in the framework of the TFE regime can be expressed as

$$J_F = J_{\mathrm{SF}}\left[\exp\left(\frac{qV}{n_F kT}\right) - 1\right] \quad \text{with} \quad n_F = \frac{E_{00}}{kT}\coth\left(\frac{E_{00}}{kT}\right) = \frac{E_0}{kT}. \tag{1.28}$$

In the reverse bias, the metal potential is raised, as shown in Figure 1.11. If the doping level in the semiconductor is low and the barrier width is large (keep in mind that the barrier width becomes smaller for energies above the Fermi level in the metal as compared to the forward-bias case), the current flow is through thermionic emission and Equation 1.7 together with Equation 1.8 treats the problem well. This process is schematically shown in Figure 1.11a. However, in cases when the doping level is moderate or high, the dominant current mechanism in the reverse-bias direction also would be TFE and FE currents, similar to the case of the forward-bias conditions as shown in Figure 1.11b. As in the case of forward bias, we can think of the TFE current being dominant in an intermediate temperature range and the FE current being dominant in low-temperature range.

Let us take up the intermediate temperature region in which we consider that when the electrons tunnel at energy E_m as defined in Figure 1.11b. With the help of

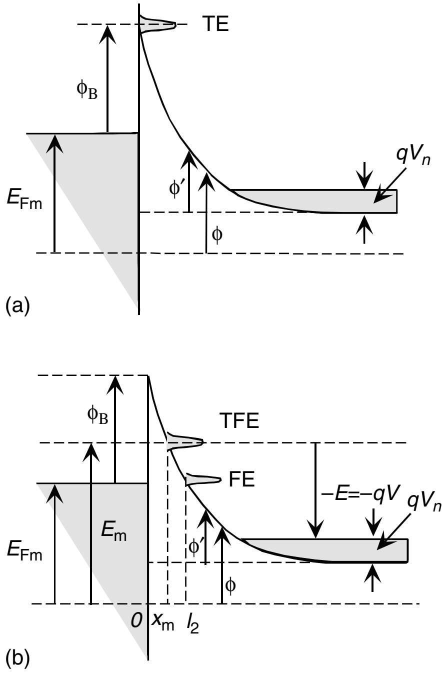


Figure 1.11 Potential energy diagram and current flow mechanisms for a reverse-biased Schottky barrier: (a) for the TE process, which is more likely when the doping level in the semiconductor is relatively low and thus the Fermi level is below the conduction band, and (b) for TFE and direct tunneling, which is also referred to as FE. The latter is more likely when the doping level in or on the semiconductor surface is sufficiently high to the extent that the Fermi level may even be in the conduction band as depicted.

Equations 1.18–1.20, the parameters $b_{\rm m}$, $c_{\rm m}$, and $f_{\rm m}$ can be obtained as

$$b_{\rm m} = \frac{q(E_{\rm m} - V - E_{\rm Fm})}{E_{00}} \left[\frac{(q\phi_{\rm B} - qV)^{1/2}(q\phi_{\rm B} + E_{\rm Fm} - E_{\rm m})^{1/2}}{E_{\rm m} - E - E_{\rm Fm}} - \log \frac{(q\phi_{\rm B} - qV)^{1/2} + (q\phi_{\rm B} + E_{\rm Fm} - E_{\rm m})^{1/2}}{(E_{\rm m} - E - E_{\rm Fm})^{1/2}} \right], \tag{1.29}$$

$$c_{\rm m} = \frac{1}{E_{00}} \log \left[\frac{(q\phi_{\rm B} - qV)^{1/2} + (q\phi_{\rm B} + E_{\rm Fm} - E_{\rm m})^{1/2}}{(E_{\rm m} - qV - E_{\rm Fm})^{1/2}} \right], \tag{1.30}$$

and

$$f_{\rm m} = -\frac{1}{4} E_{00} \left[qV - \frac{q\phi_{\rm B}}{\cosh^2(E_{00}/kT)} \right], \tag{1.31}$$

where the energy $E = qV$, referring to Figure 1.11b, with V being the amplitude of the applied reverse bias.

Using Equations 1.17 and 1.31, one can now derive an expression for E_m, the energy at which tunneling occurs, as

$$E_m = E_{Fm} + \frac{q\phi_B - qV\sinh^2(E_{00}/kT)}{\cosh^2(E_{00}/kT)}. \tag{1.32}$$

The current–voltage relationship in this has again an exponential dependence that can be expressed as

$$J_{tfeR} = J_{SR}\exp\left(\frac{-qV}{E'}\right), \tag{1.33}$$

where

$$E' = E_{00}\left[\frac{E_{00}}{kT} - \tanh\left(\frac{E_{00}}{kT}\right)\right]^{-1} \tag{1.34}$$

and

$$J_{tfes} \equiv J_{SR} = \frac{AT^2\sqrt{\pi q E_{00}(\phi_B - V\cosh^2(E_{00}/kT))}}{kT\cosh(E_{00}/kT)}\exp\left(-\frac{q\phi_B}{E_0}\right), \tag{1.35}$$

where A is now the Richardson constant of the metal.

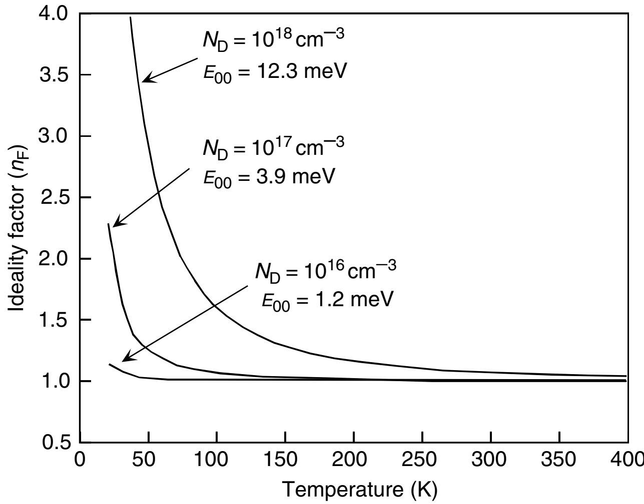
Considering the electron emission from the semiconductor to the metal, the reverse current–voltage characteristics for the thermionic field emission region can be expressed in terms of more familiar parameters such as

$$J_R = J_{SR}\left[\exp\left(\frac{qV}{n_R kT}\right) - 1\right] \quad \text{with} \quad n_R = \frac{E_{00}}{kT}\left[\frac{E_{00}}{kT} - \tanh\left(\frac{E_{00}}{kT}\right)\right]^{-1}. \tag{1.36}$$

The term J_{ss} represents the saturation current as in the case of TE but with very different functional dependence. These relations provide the smooth transition from the TFE regime to just FE regime as the temperature is lowered, which hampers the thermionic emission. A unique property here is that the sum of the inverse of the forward and reverse ideality factors adds up to 1.

$$n_F^{-1} + n_R^{-1} = 1. \tag{1.37}$$

The temperature dependence of the ideality factor calculated for three doping levels, namely, 10^{16}, 10^{17}, and $10^{18}\,\text{cm}^{-3}$ with the aid of Equations 1.28 and 1.36 reported in Ref. [8] are shown in Figure 1.12 along with the E_{00} parameter. The ideality factor measured as a function of temperature for Pt and Ni Schottky barriers along



1. $N_D = 10^{18}$cm^{-3} =10^{-24}m^{-3}, E_{00} = 12.3 meV
n_F (T = 50 K) = 2.8676
n_F (T = 300 K) = 1.0740
n_F (T = 400 K) = 1.0419
2. $N_D = 10^{17}$cm^{-3} =10^{-23}m^{-3}, E_{00} = 3.9 meV
n_F (T = 50 K) = 1.2569
n_F (T = 300 K) = 1.0075
n_F (T = 400 K) = 1.0042
3. $N_D = 10^{16}$cm^{-3} =10^{-22}m^{-3}, E_{00} = 1.2 meV
n_F (T = 50 K) = 1.0269
n_F (T = 300 K) = 1.0008
n_F (T = 400 K) = 1.0004

Figure 1.12 Ideality factor versus temperature for three doping levels, namely, along with the corresponding E_{00} parameters in the temperature range of about 10–20 to 400 K.

with the calculated values for E_{00} parameter in the range of 13–22 meV showed the E_{00} parameter to be 20 meV for Pt and between 16 and 17 (assumed to be 17 from here onward) meV for Ni Schottky contact as shown in Figure 1.13.

The measured temperature-dependent I–V characteristics for both forward- and reverse-bias directions are shown in Figure 1.14a and b when the sample surface was cleaned in NaOH prior to metallization. In addition, the forward characteristics when the sample was cleaned in buffered HF are shown in Figure 1.14c with anomalous plateaus. Although not observable in all samples, the data are consistent with some of the reports in the literature [9,10]. Clearly, the data cannot be modeled with TE model only, which gives only a simple exponential behavior. However, the measured values agree well if one uses the TEF and TE models with one complication. The data and calculations agree if one assumes $E_{00} = 17$ meV for the Ni/GaN contact and $E_{00} = 20$ meV for the Pt/GaN contact in Equations 1.28 and 1.36, respectively. However, the calculated value for E_{00} using Equation 1.21 and a bulk doping

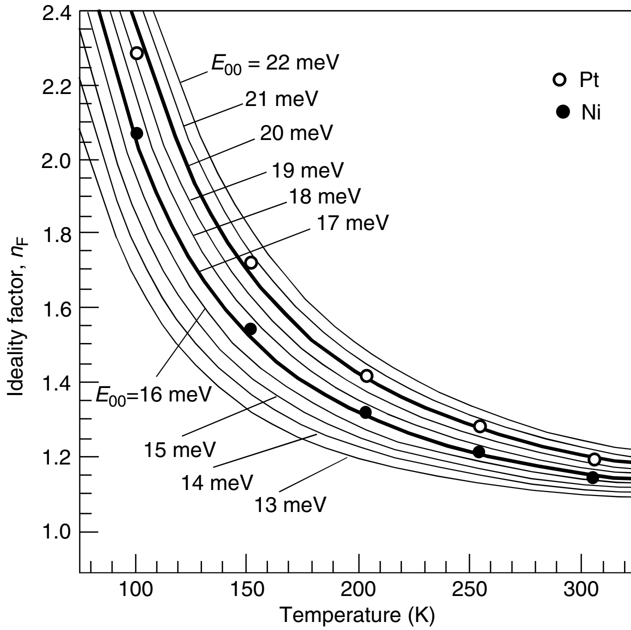


Figure 1.13 The measured ideality factors for Pt (open circles) and Ni (full circles) Schottky barriers along with the prediction from the TFE model given in Equation 1.28 using E_{00} parametric values in the range of 13–22 meV [8].

$N_D = 1.0 \times 10^{17}\,cm^{-3}$ (the measured value) is 4 meV, which is much smaller than those of Figure 1.13. To obtain 17 and 20 meV for E_{00}, doping level has to be increased to 2.2×10^{18} and $3.1 \times 10^{18}\,cm^{-3}$, respectively. These values are too high to be attributed to doping fluctuations in the bulk donor concentration. Moreover, without the sum of the inverse values of the ideality factors in forward and reverse directions being equal to 1, Equation 1.37 is violated. Furthermore, the TFE/TE model does not predict the measured *I–V* data quantitatively. This simply means that the effect of surface and bulk defects must be considered. As will be shown below, a better agreement with experiments is obtained when a higher doping concentration is assumed near the surface as compared to the bulk. One must remember this is simply a model in that while a higher doping concentration near the surface may allow one to fit the *I–V* characteristics better, the root cause of the problem most likely lies in the defects in the material and anomalously high current may be because of the sequential tunneling mitigated by these defects.

To underscore the dramatic effect of the E_{00} parameter on the *I–V* characteristics, calculations of Hasegawa *et al.* [8] for $E_{00} = 4$ meV and $E_{00} = 17$ meV are shown in Figure 1.15 for both forward- and reverse-bias conditions with temperature as a parameter. The 17 meV value increases the reverse current by orders of magnitude in addition to initiating substantial current flow at very small reverse-bias conditions. The same trend holds for forward characteristics as well, albeit not to the same extent as shown in Figure 1.15.

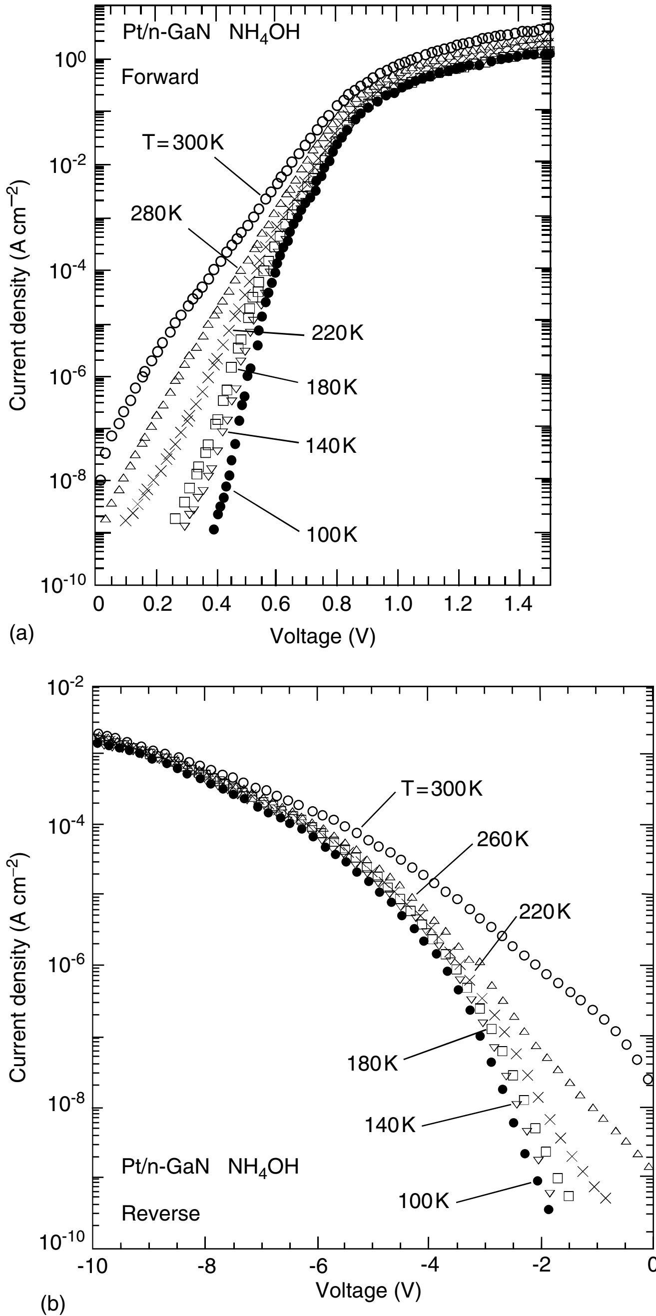


Figure 1.14 Measured (a) forward and (b) reverse *I–V–T* characteristics of a Pt/GaN Schottky contact. (c) Forward *I–V–T* characteristics showing anomalous current plateaus when the surface is cleaned in buffered HF [8].

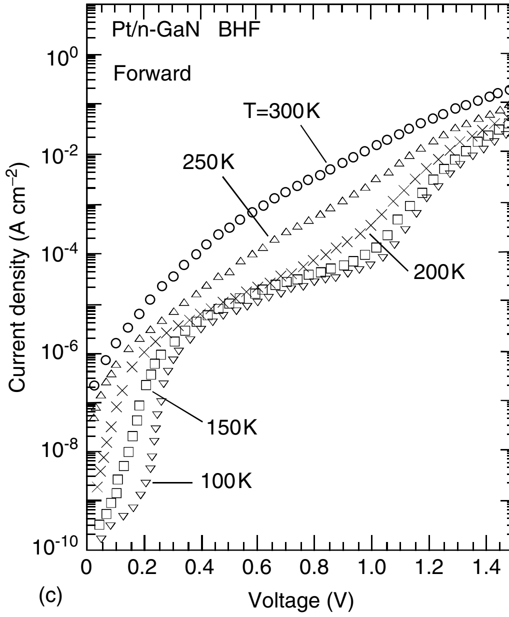


Figure 1.14 (*Continued*)

1.2.3
Direct Tunneling Regime

At low temperatures and high doping concentrations, direct tunneling, FE, process dominates the current flow in a metal–semiconductor contact. Referring to Figure 1.10b for the forward and Figure 1.11 for the reverse-bias conditions, and also Figure 1.6, the density of current flowing from the semiconductor to the metal is proportional to the product of the transmission coefficient, the occupation probability in the semiconductor, f_s, and the unoccupation probability in the metal, $1-f_m$,[1)]

$$J_{s\to m} = \frac{A^*T}{k}\int_0^{q\phi_B} f_s T(\xi)(1-f_m)\,d\xi, \tag{1.38}$$

where $T(\xi)$ is the transmission coefficient and is given by, for low temperatures and/or high doping levels, $T(\xi) \approx \exp(-q\phi_B/E_{00})$. Similarly, the density of current flowing from the metal to the semiconductor is proportional to the product of the transmission coefficient, the unoccupation probability in the semiconductor, and the occupation probability in the metal is

$$J_{m\to s} = -\frac{A^*T}{k}\int_0^{q\phi_B} f_m T(\xi)(1-f_s)\,d\xi. \tag{1.39}$$

1) The occupation probability depicts the likelihood that a state *is* occupied by an electron, and one minus the occupation probability exhibits that to be free of electrons.

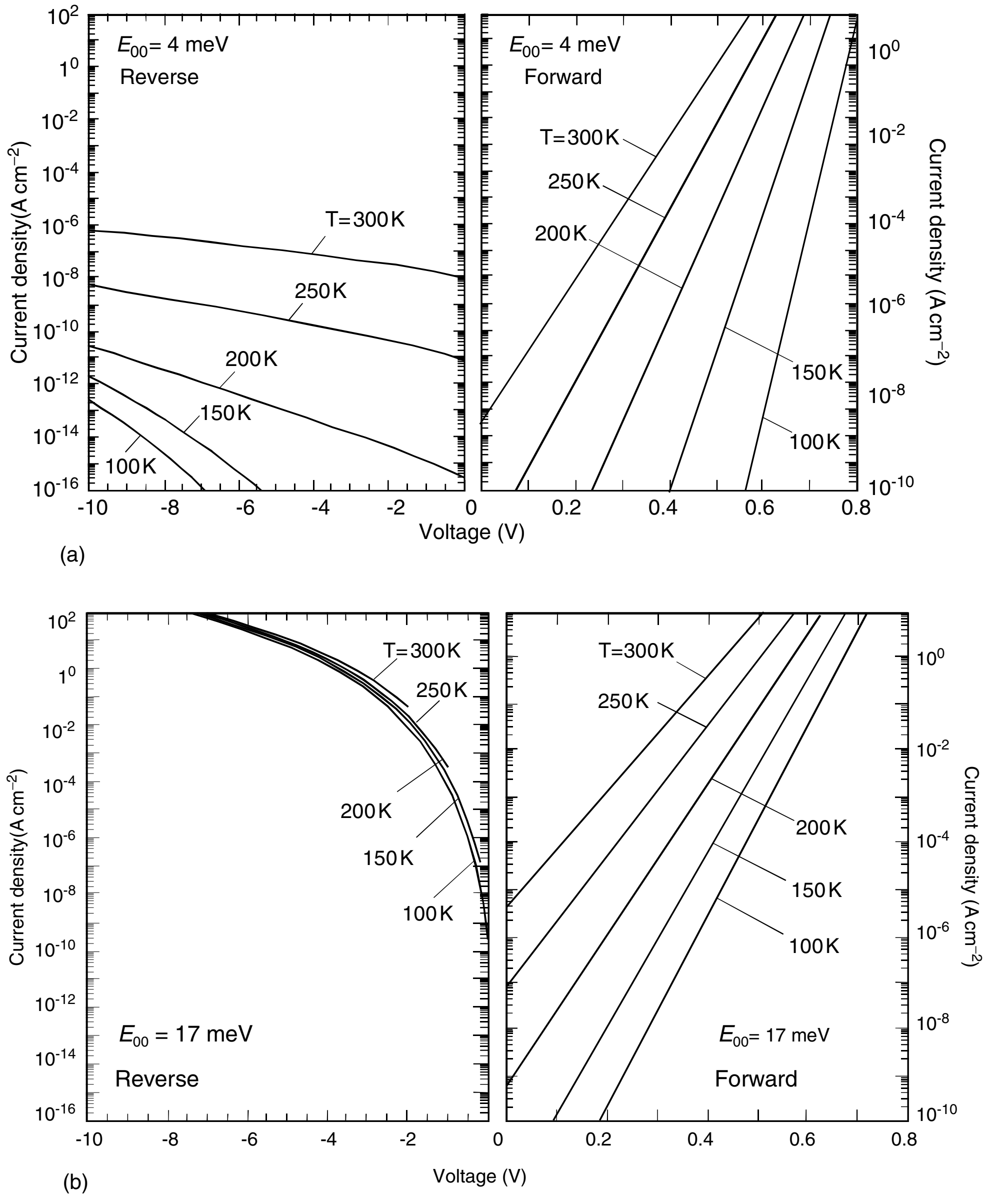


Figure 1.15 Calculated *I–V–T* curves using the TFE/FE model (a) for $E_{00} = 4$ meV, which corresponds to the bulk doping of the sample, and (b) for $E_{00} = 17$ meV obtained in Figure 1.13 for the Ni/GaN contact [8].

The total density of current that is simply the sum of the density of current flowing in both directions can be approximated by

$$J_t \approx \exp\left(\frac{-q\phi_B}{E_{00}}\right). \tag{1.40}$$

For the forward-bias case, the parameters b_m, c_m, and f_m in Equation 1.16 but now for the field emission current are defined as

$$b_m = \frac{q\phi_B - qV}{E_{00}}, \tag{1.41}$$

$$c_m = \frac{1}{2E_{00}}\log\left(\frac{4(\phi_B - V)}{V_n}\right), \tag{1.42}$$

and

$$f_m = \frac{1}{4}E_{00}qV_n. \tag{1.43}$$

Padovani and Stratton [7] presented an analytical expression for the forward current for direct tunneling as

$$J_{FE} = J_{SFE}\exp\left(\frac{qV}{E_{00}}\right), \tag{1.44}$$

and

$$J_{SFE} = \frac{2\pi A^* T^2 E_{00}}{kT\left[\log\left\{2\left(\frac{(\phi_B - V)}{V_n}\right)\right\}\right]\sin\left[\frac{\pi kT}{2E_{00}}\log\left\{2\left(\frac{\phi_B - V}{V_n}\right)\right\}\right]}\exp\left(\frac{-q\phi_B}{kT}\right). \tag{1.45}$$

For the reverse-bias case, the parameters b_m, c_m, and f_m are defined as

$$b_m = \frac{1}{E_{00}}\left[\sqrt{q\phi_B(q\phi_B - qV)} + qV\log\frac{\sqrt{q\phi_B} + \sqrt{q\phi_B - qV}}{\sqrt{-qV}}\right], \tag{1.46}$$

$$c_m = \frac{1}{E_{00}}\log\frac{\sqrt{q\phi_B} + \sqrt{q\phi_B - qV}}{\sqrt{-qV}}, \tag{1.47}$$

and

$$f_m = -\frac{1}{4}E_{00}E. \tag{1.48}$$

The above parameters, however, are not easily traceable in terms of experiments. By assuming that the reverse-bias voltage is larger than the barrier height, Padovani and Stratton further simplified Equations 1.45–1.47 to

$$b_m = \frac{2\phi_B^{3/2}}{3E_{00}\sqrt{q\phi_B - qV}} \quad \text{and} \quad c_m = \frac{\phi_B^{1/2}}{E_{00}\sqrt{q\phi_B - qV}}. \tag{1.49}$$

The resultant *I–V* characteristic is then

$$J_{\mathrm{FER}} = \frac{\pi A E_{00} T^2}{kT\left[\sqrt{\frac{\phi_B}{\phi_B - V}}\right]\sin\left\{\frac{\pi kT}{E_{00}}\sqrt{\frac{\phi_B}{\phi_B - V}}\right\}}\exp\left(\frac{-2(q\phi_B)^{3/2}}{3E_{00}\sqrt{q\phi_B - qV}}\right), \tag{1.50}$$

where A is the Richardson constant of the metal. In the limit of zero temperature, Equation 1.50 further reduces to

$$J_{\mathrm{FER}} = AT^2\left(\frac{E_{00}}{kT}\right)^2\frac{\phi_B - V}{\phi_B}\exp\left(\frac{-2(q\phi_B)^{3/2}}{3E_{00}\sqrt{q\phi_B - qV}}\right). \tag{1.51}$$

Equation 1.51 shows that a plot of ln $J_{\mathrm{FER}}/q(\phi_B - V)$ as a function of $\sqrt{q\phi_B - qV}$ would yield a straight line with a slope of $2(\phi_B)^{3/2}/3E_{00}$.

A plot of the ln of the current versus voltage given in Equation 1.44 would yield a slope of q/E_{00} regardless of the temperature that is a characteristic of direct tunneling current. The coefficient in front of Equation 1.44 (the saturation current), which is expanded in Equation 1.45, clearly indicates that the lower the barrier, ϕ_B, and the higher the doping level, which increases E_{00} through Equation 1.21, the higher the saturation current and thus the higher the current. This explicitly implies that the resistance is low. The key for a good ohmic contact and high current is then to find a metal with a small barrier, which calls for a metal with a work function equal to or smaller than that of the semiconductor for the n-type case. For the p-type case, the same implies in that the work function of the metal needs to be equal to or larger than that of the semiconductor, which is very hard to do for large-bandgap semiconductors. Additionally, the situation is exacerbated by the large hole mass, which tends to decrease the E_{00} term and thus increase the resistance. In fact, the low hole concentration combined with a large hole effective mass precludes direct tunneling for all practical purposes in p-type GaN. What may be happening in the presently utilized Ni/Au contacts is that (to be discussed in Section 1.6), Ni/Au chemically modifies the GaN surface and the current conduction mechanism may be dominated by defect-assisted tunneling, which we will call *leakage current* (Section 1.2.4). This is nearly impossible to model as the nature of the defects germane to this problem is not known. However, models to lump defect-related processes into a surface layer or a high concentration of defects near the surface could be used to fit the current–voltage characteristics as shown in Sections 1.2.2 and 1.2.3. Naturally, as materials quality improves, so does the fitting of the current–voltage characteristics without causing drastic defect-induced or surface-induced effects. With improvements in the quality of p–n or metal–semiconductor junctions, one can extract the leakage component from temperature-dependent current measurements with the help of the current conduction models discussed in this subsection. This has been undertaken for Schottky barriers in n-type GaN [11].

1.2.4 Leakage Current

In addition to thermionic emission, thermionic field emission, and tunneling currents, other currents such as defect-assisted tunneling, which may have a quasiohmic nature, are typically lumped in the leakage component of more or less unknown origin and can be expressed as

$$I_{\text{lk}} = \frac{V - IR_{\text{s}}}{R_{\text{t}}}, \tag{1.52}$$

where R_{t} is considered to be a fitting parameter, which represents defects and inhomogeneities at the metal–semiconductor interface. In semiconductors with less than ideal interfaces, a tunneling barrier E_0 may not be predicted, in which case it should be considered as another fitting parameter. In practice, the terms $I_{\text{te}\,0}$ and $I_{\text{tfe}0}$ are also considered to be fitting parameters that represent the magnitudes of the contributions to the current from thermionic emission and thermionic field emission, respectively.

While the above discussion does represent a Schottky barrier contact on nearly defect-free GaN, the GaN is anything but nearly defect free. Thus, the current–voltage characteristics deviate substantially from the aforementioned treatment and the role that surfaces and defects play must be taken into consideration. This has a bearing not only on the I–V characteristics in both forward- and reverse-bias directions but could also be the dominating component in reverse bias. This takes on a special meaning in Schottky detectors where the reverse-bias leakage goes toward the dark current and adversely affects the detectivity of the device as discussed in Volume 3, Chapter 4.

In practice, the current–voltage characteristic is depicted by

$$J = J_0\left[\exp\left(\frac{qV}{nkT}\right) - 1\right], \tag{1.53}$$

when dominated by thermionic emission with n being the ideality factor and lumping deviations from ideal thermionic emission. The J_0 term, which is also designated by J_{s} in many textbooks and papers, is the saturation current for ideal thermionic emission as given by Equation 1.8 and is to a first extent independent of voltage except through any barrier lowering. However, the I–V characteristics of GaN Schottky barriers cannot simply be modeled by ideal thermionic emission owing to bulk and surface defects.

Assuming that defects do play a role, which is elaborated on in Section 1.2, and to further reduce low bias dark or leakage currents and noise power, better insight is required into the defect-assisted tunneling frequently encountered. The tunneling current also coincides with the rapid degradation of the current–voltage characteristics under voltage stress. The strong voltage dependence of the dark current at low bias is indicative of tunneling and was investigated by Carrano *et al.* [12,13] who suggested a series of sequential deep-level assisted processes. A schematic band diagram of the metal–semiconductor interface with defects likely responsible for tunneling is shown in Figure 1.16. In this scenario, following the initial application of

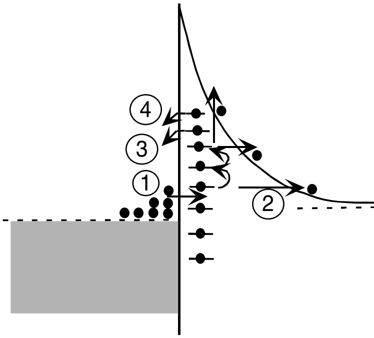


Figure 1.16 Band diagram of a Schottky diode interface illustrating specific defect-assisted tunneling processes proposed for explaining the large leakage at low bias voltages [12].

a small external bias, electrons tunnel from the Schottky barrier to an interfacial state labeled (1) in Figure 1.16. At this point, two or more paths become available, namely, tunneling through the remaining barrier, which is depicted as process (2), or thermal excitation through a set of deep states, which is depicted as process (3). Participation by defects introduces memory effects owing to time constants involved, in that trap filling and ionization with their characteristic time constant are involved. This manifests itself as unstable current–voltage characteristics as the voltage across the device is swept. For example, the completion of the first voltage sweep would cause for most of the available defect states to be filled. This leaves just a few empty states available for deep-level assisted tunneling. Therefore, the second voltage sweep would show a current–voltage characteristic with lower leakage current, albeit unstable. However, once the filled state population reaches equilibrium, the current–voltage characteristic becomes stable. It is also plausible to release trapped electrons, depicted as process (4) [8], by tunneling back to the Schottky metal. This would give rise to excess leakage current at low bias voltages. Some of the detrapping processes appear to be strongly influenced by illumination with white light. Light excitation frees defect states and results in increased leakage current when dark current measurements are made following illumination in Schottky barrier detector structures. The foregoing discussion would not be needed had it not been for the relatively poor quality of GaN. It is therefore expected that as the quality of GaN is increased, the anomalies in the Schottky barrier characteristics would be minimized if not fully eliminated. With nano-ELO using SiN_x nanonet, Xie *et al.* [14] produced GaN layers with a total dislocation density of mid $10^7\,cm^{-2}$ resulting in nearly ideal Schottky barriers. For the most part, the current–voltage characteristics could be accounted for thermionic emission. In the interim thought, models, not necessarily theories, have been developed to fit the experimentally observed I–V characteristics with theory by invoking a model highly conductive surface layer the source of which has been attributed to processing. Again, this model designed to fit the experiments to the theory should be treated as such.

As mentioned above, a plausible explanation for anomalous current transport, trap-assisted tunneling transport is invoked. The presence of discrete deep-level states or a continuum of trap-induced states near the surface next to the Schottky barrier is assumed to provide tunneling paths through the energy barrier after the capture of electrons by thermal excitation. However, in terms of modeling purposes, it is difficult to prescribe the defect levels, defect energy distribution, and defect concentration accurately. However, the introduction of a thin surface layer mimicking a highly doped layer would enhance the current bringing the calculated and observed currents closer [8]. This model is schematically shown in Figure 1.17 with the assumed presence of the thin surface barrier (TSB) regions having a thickness d. Given the nonuniformity and disparity of the GaN layers and surfaces which also involve high densities of dislocations, it is very likely that the exact nature of the TSB region would be sample and device dependent, the latter representing the spatial nonuniformity.

For a quantitative description, Hasegawa *et al.* [8] also assumed that lateral extent of the thin surface layer is much larger than the value of d, allowing for a one-dimensional treatment of the current transport. In Figure 1.17, the TSB region is characterized by a net surface donor concentration of N_{DS}, much higher than the bulk value, a thickness d, and an effective barrier height for the TFE/FE process ϕ_{B}, which includes the image force barrier lowering. The potential at the boundary $x = d$ is defined as ϕ_d. The potential shape in the TSB region is a sharp parabola, whose minimum potential is defined as ϕ_0. The fitting to be discussed indicates that the sheet concentration of ionized donors given by the $N_{\mathrm{DS}}d$ product is in the vicinity of $10^{12}\,\mathrm{cm}^{-2}$, which seems to be a very reasonable number of surface defects.

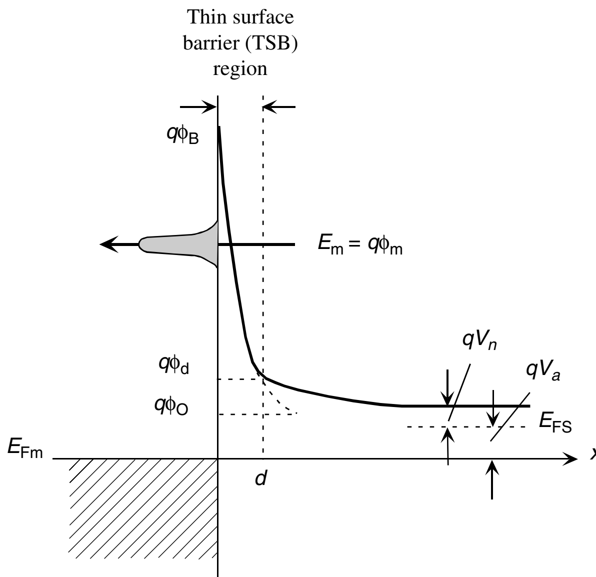


Figure 1.17 Thin surface barrier model in the form of a schematic band diagram as proposed in Ref. [8].

Let us define V_0 as the bias voltage at which $\phi_0 = \phi_d$ holds, then V_0 and ϕ_0 are given by the following equation:

$$V_0 = \phi_B - \frac{qN_{DS}}{2\varepsilon_s} d^2 - V_n. \tag{1.54}$$

For $V > V_0$,

$$\phi_0 = \frac{qN_{DS}}{2\varepsilon_s}\left(1 - \frac{N_D}{N_{DS}}\right)\left[\sqrt{\frac{2\varepsilon_s}{qN_D}(V_0 - V) + d^2} - d\right]^2 + V + V_n. \tag{1.55}$$

For $V < V_0$,

$$\phi_0 = V + V_n. \tag{1.56}$$

Following the steps of Padovani and Stratton [7], one can calculate current owing to the TFE/FE process. It has turned out that the value of the minimum potential of the TSB parabola ϕ_0 is an important controlling parameter for the TFE/FE processes through the TSB region, although this minimum is located outside the TSB region and does not appear in the real potential profile for $V < V_0$. In fact, the energy location of the Gaussian main peak $q\phi_m$ for the TFE process is given by

$$\frac{\phi_m - \phi_0}{\phi_B - \phi_0} = \frac{1}{\cosh^2(E_{00}/kT)}, \tag{1.57}$$

even for $V < V_0$ as long as the tunneling process takes place through the TSB parabolic barrier.

It can be shown that the forward and reverse currents are again given by Equations 1.28 and 1.36. However, V_a must be replaced with ϕ_0 and the value of n_F as described in Equation 1.28 must be determined using the surface donor concentration N_{DS}. The average forward and reverse saturation current densities in the TFE regimes are given by Equations 1.27 and 1.35 which after the aforementioned modification take the form

$$J_{0F} = A^* T^2 \frac{\sqrt{\pi E_{00} q(\phi_B - \phi_0 - V_n)}}{kT\cosh(E_{00}/kT)} \exp\left(-\frac{qV_n}{kT}\right)\exp\left(\frac{-q(\phi_B - V_n)}{n_F kT}\right), \tag{1.58}$$

$$J_R = A^* T^2 \frac{\sqrt{\pi E_{00}}}{kT}\left[qV_R + \frac{q\phi_B}{\cosh^2(E_{00}/kT)}\right]^{1/2} \exp\left[\frac{-q\phi_B}{n_F kT} + \frac{qV_R}{n_R kT}\right] \quad \text{for TFE}, \tag{1.59}$$

$$J_R = A^* T^2 \frac{\pi E_{00}\exp[-2q^{3/2}\phi_B^{3/2}/3E_{00}(q\phi_B - q\phi_0)^{1/2}]}{kT[q\phi_B/(q\phi_B - q\phi_0)]^{1/2}\sin(\pi kT[q\phi_B/(q\phi_B - q\phi_0)]^{1/2}/E_{00})} \quad \text{for FE}. \tag{1.60}$$

It can be shown that at high reverse voltages, the reverse ideality factor satisfies the following relation:

$$n_R = \frac{N_{DS}}{N_D} \frac{1}{1 - 1/n_F}. \tag{1.61}$$

This equation is consistent with Equations 1.28 and 1.36 once different doping levels in the bulk and on the surface are taken into consideration.

To compare the prediction of the TSB model with the measured data shown in Figure 1.14, efforts were made to fit the theoretical *I–V–T* curves to experimental data [8]. The results are given in Figure 1.18 for the Pt/GaN contact. In the fitting, best values of E_{00} and ϕ_B to reproduce the forward *I–V–T* curves were chosen first for the bulk doping of $N_D = 1.0 \times 10^{17}\,\text{cm}^{-3}$ and then, the value of d was determined to give the best fit to the reverse *I–V–T* curves. The parameter values used in Figure 1.18 were $E_{00} = 20\,\text{meV}$, $N_{DS} = 3.3 \times 10^{18}\,\text{cm}^{-3}$, $\phi_B = 1.05\,\text{eV}$, and $d = 23.5\,\text{nm}$ for the Pt/GaN contact. The same parameters for Ni contact giving the best fit are $E_{00} = 17\,\text{meV}$, $N_{DS} = 2.1 \times 10^{18}\,\text{cm}^{-3}$, $\phi_B = 0.90\,\text{eV}$, and $d = 23\,\text{nm}$ for the Ni/GaN contact, the figures for which are not shown but can be found in Ref. [8]. For the reverse current, Equation 1.59 was used in a low bias region and Equation 1.60 was used in a high bias region.

Despite improved fitting when the highly doped surface layer model is invoked, there is still discrepancy at low temperatures and low reverse voltages region where TFE currents are extremely sensitive to the detailed shape of the barrier. Because the current conduction is exponential of the barrier in the TFE/FE process, the calculated *I–V–T* characteristics are very sensitive to even a slight change in the TSB parameter values. This model also does not account for the plateaus shown in Figure 1.14, but they can be modeled qualitatively by the band diagram shown in Figure 1.19, although the actual picture must have to do with defects and ensuing defect-assisted sequential tunneling. Region 1 in Figure 1.19 represents the case of $V > V_0$ where the depletion layer is within the TSB layer and the TFE is directly controlled by V leading to a straight $\log J_F$ versus V line with a slope given by Equation 1.28. Region 2 is for the $V_0 > V > V_D$ case where V_D is defined as the voltage at which $\phi_d = \phi_m$ meaning that the peak energy of TFE shown as a beam in the figure, coincides with the boundary energy of two parabolas for TSB and bulk band potential profiles. In this region, the TFE process takes place through the TSB region, but its dependence on voltage is reduced owing to the depletion region extending below the TSB region, which causes a partial voltage drop in the bulk region, giving rise to nearly a plateau. Finally, region 3 is for $V_D > V > 0$, which represents the case where TFE starts to take place partially through the bulk layer and thinner barrier. The occurrence of this plateau has also been attributed to a lateral in homogeneity of the Schottky barrier height itself. Again, if this is so one has to address the question for the root cause of Schottky barrier nonuniformity, which again would have its genesis in the defects.

Consequently, for a complete quantitative analysis, a good understanding of the root causes is imperative. Once that is available, a full numerical analysis of the TFE/FE process, taking into account various potential profiles, Schottky barrier heights, and nonuniformities, would be needed to more accurately model the current–voltage

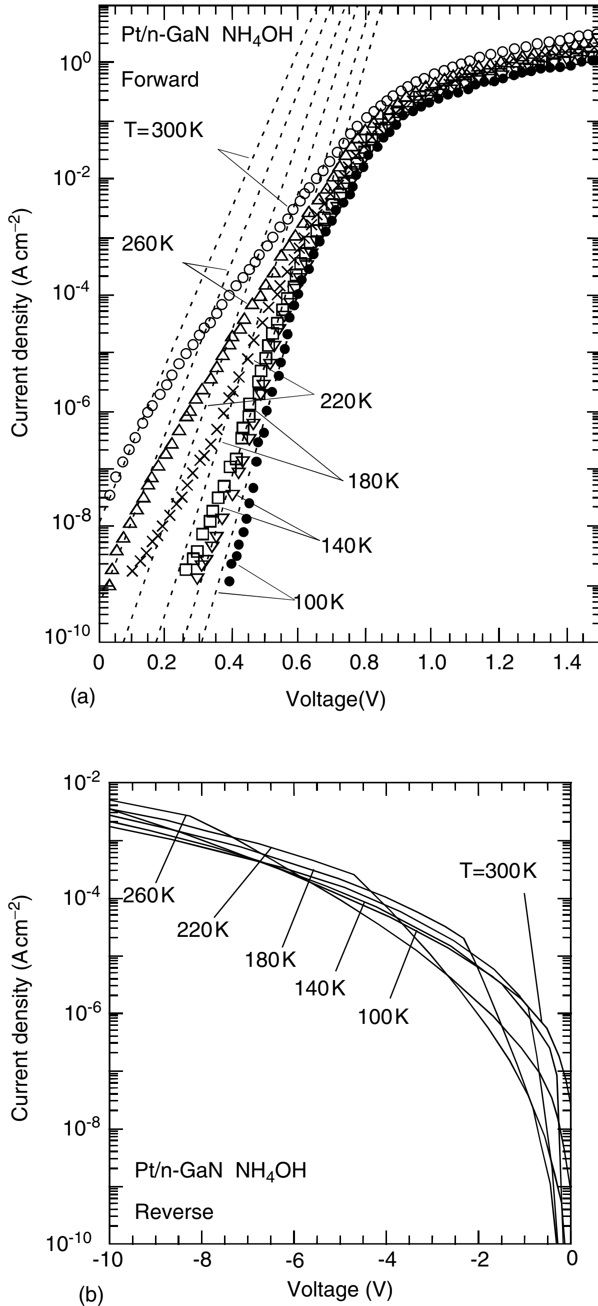


Figure 1.18 Calculated (a) forward and (b) reverse *I–V–T* characteristics for Pt/GaN contact. The dashed lines in (a) represent the thermionic emission model only [8].

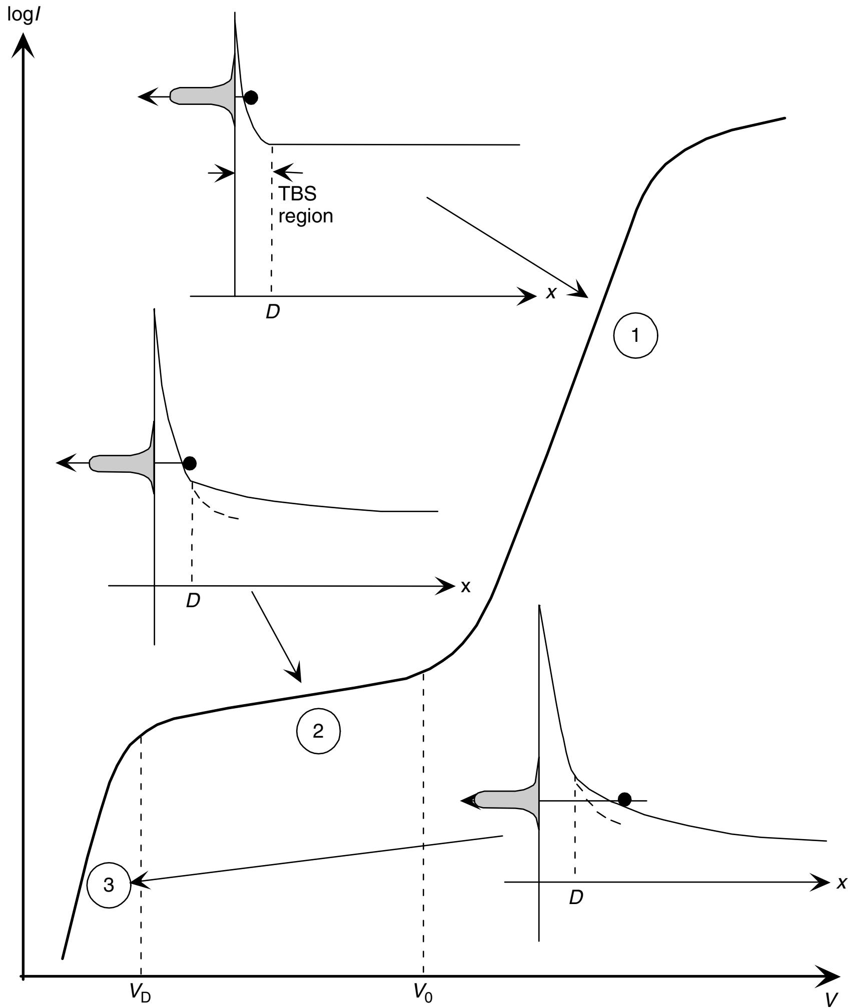


Figure 1.19 Schematic band diagram facilitating the qualitative description of plateaus observed in the forward *I–V* characteristics of some of the devices [8].

characteristics of Schottky barriers on GaN. It should also be mentioned that the surface of GaN is not completely inert [15], and any trap whose energy lies between the conduction band of the n-type semiconductor and the Fermi level in the metal would participate in this process.

An additional anomaly in some Ni and Pt GaN Schottky contact samples is that the samples show hysteresis at low temperatures, indicating participation of trap states. Owing to this behavior, the reverse *I–V–T* data given previously were taken along the direction of reducing the bias that was always much more reproducible. This is most

probably because of the fact that a dynamic equilibrium of trap occupancy is much more easily established under a flow of high leakage currents.

This would make a good segue to discuss the effect of defects on I–V characteristics. The state of semiconductor science is such that it is not yet truly possible to determine directly if there are point defects under a Schottky barrier and if so in what form. However, it is reasonable to assume, and there is some evidence for this, which is discussed in Volume 1, Chapter 4, that regions of the semiconductor containing extended defects also have point defects. If so, correlation between the I–V performance and the presence or absence of extended defects could be made, as has been attempted by Shiojima and Suemitsu [16]. The authors used a combination of submicrometer-sized Schottky dot array on the surface and conductive atomic force microscopy (CAFM). The latter was used to determine if the observable dislocations contribute to extraneous current conduction paths. The conclusion is that mixed and edge dislocations are inactive as they did not affect the Schottky barrier height, ideality factor, and reverse leakage current. However, large structural defects with a diameter of a few hundred nanometers shorted the Schottky contacts. The density of these large structural defects, however, was less than $4 \times 10^5\,\mathrm{cm}^{-2}$. This seemingly contradictory conclusion to the prologue of the paragraph calls for further investigations to determine the root causes of anomalies observed in I–V characteristics.

Other current conduction mechanisms, particularly those observed under reverse bias, in Schottky barrier contacts are in some ways very similar to those in p–n junctions. Among atypical ones, which are not commonly observed in high-quality semiconductors, are hopping conduction, Poole–Frenkel current, and perhaps tunneling current. Experimental results, particularly in p–n junctions, indicate that the reverse current exponentially depends on temperature eliminating the possibility of tunneling current. This leaves behind the effects of hopping conduction and Poole–Frenkel current conduction, both of which are discussed in Section 4.4.4. When the above-mentioned current components are small enough, the reverse-bias voltage can be increased without burning the device and at a critical field avalanching occurs. At this stage, the carriers in depletion region would be energetic enough to lose their energy through the process of creating electron–hole pairs. This can carry on to create large numbers of carriers participated by not only the primaries, but also the secondaries, and so on. This process is called avalanche multiplication and is reversible if measures are taken to thermally protect the device. The process has been successfully used in some two terminal microwave devices as well as photodetectors with internal gain. The mechanism is discussed in more detail in Section 4.4.5, in conjunction with p–n junctions.

For completeness, in cases of highly doped semiconductors and surface oxide layers in the context of metal–semiconductors, there are other current conduction mechanisms such as Fowler–Nordheim. Field emission with the aid of thermionic emission and tunneling through a triangular barrier is called the Fowler–Nordheim tunneling – wherein the triangular barrier is caused by the application of an external field [17,18]. (Note that a correction to the $\nu(y)$ function appearing in the probability calculations in the original paper has been corrected in a later publication by

Burgess *et al.* [19].) The process is of prominence in field emission from metal into vacuum where there is electric field, in metal oxide structures where the oxide band is triangular-like owing to applied bias, and metal–semiconductor cases where the semiconductor is heavily doped. The metal oxide case is very pertinent to nonvolatile memories with gates buried in SiO_2. As mentioned, the process was first described by Fowler and Nordheim, thus the namesake, in conjunction with emission from a metal into vacuum, where there is an electric field. As such, the mechanism is naturally applied to electron emission from metal tips, carbon nanotubes, and so on to vacuum, in addition to the context relevant to semiconductors.

The tunneling probability can be derived using the time-independent Schrödinger equation, the solution of which within the Wigner, Kramer, Brillouin (WKB) approximation would lead to tunneling probability for a triangular barrier [20–22]

$$P = \exp\left[\frac{8\pi}{3hE}\sqrt{2qm^{*}\phi_{B}^{3}}\right]. \tag{1.62}$$

The tunneling current can then be found by taking the product of the velocity, density, tunneling probability, and the electronic charge as follows:

$$J_{FN} = qnvP = qnv\exp\left[\frac{8\pi}{3hE}\sqrt{2qm^{*}\phi_{B}^{3}}\right]. \tag{1.63}$$

The tunneling current, therefore, depends on exponentially to the $3/2$ power of the barrier height and most characteristically the ln of current divided by the field square, as will be seen soon, is dependent linearly on the $1/E$. When the missing details are introduced, the Fowler–Nordheim tunneling current is fully expressed as [21]

$$J_{FN} = C_{FN}E^{2}e^{(-B/E)}, \tag{1.64}$$

where E is the electric field at the metal insulator interface (the insulator here is either an inadvertent oxide layer formed during processing and/or storage or some high-resistivity layer that might be present near the surface). C_{FN} and B are constants that depend on the properties of the metal and insulator barrier height and the effective mass of electrons in the insulating layer. For the standard tunneling, these constants are given by

$$C_{FN} = \frac{q^{3}}{8\pi hq\phi_{B}} = 1.541\times10^{-6}\frac{1}{q\phi_{B}} \quad (\mathrm{A\,V^{-2}}),$$

$$B = \frac{8\pi}{3qh}\sqrt{2m_{ins}^{*}(q\phi_{B})^{3}} = 6.83\times10^{7}\sqrt{\frac{m_{ins}^{*}}{m_{e}}}(q\phi_{B})^{3/2} \quad (\mathrm{V\,cm^{-1}}), \tag{1.65}$$

where q is the electronic charge, h is the Planck's constant, m_{ins}^{*} is the electron effective mass in the insulator, and m_e is the free electron mass. Note that ϕ_B here is taken in terms of volts for consistency (in the literature, it is also expressed to be in terms of electron volts, which is why we have $q\phi_B$ in place of ϕ_B in the literature). If it is taken as energy, the $q\phi_B$ terms should be replaced with ϕ_B only. These expressions were originally developed for tunneling from metals into vacuum. Therefore, in addition to the applications in metal insulator contacts,

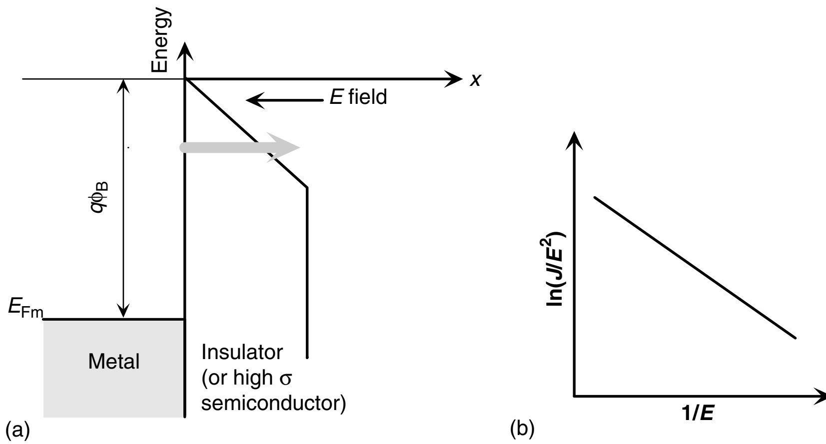


Figure 1.20 (a) Graphical description of the triangular potential barrier setting up the Fowler–Nordheim tunneling; (b) current–voltage relationship in the form of $ln(J/E^2)$ versus $1/E$. Knowing the pertinent parameters, the barrier height can be determined from this plot. In the case of quantum mechanical reflection requiring smooth and sharp interfaces, the $ln(J/E^2)$ versus $1/E$ relationship assumes an oscillatory behavior around the linear dependence shown.

the concept is also applicable to emission of electrons from any tip into vacuum, such as carbon nanotubes, semiconductor, or metal field emission tips. Moreover, the electronic process described here is directly applicable to floating gate nonvolatile memories. A graphical description of the triangular potential barrier setting up the Fowler–Nordheim tunneling as well as current–voltage relationship in the form of $\ln(J/E^2)$ versus $1/E$ is shown in Figure 1.20. Quantum mechanical reflections from interfaces impart an oscillatory behavior on top of the linear $\ln(J/E^2)$ versus $1/E$ relationship providing that the interfaces are sharp and smooth. It should be mentioned that other current conduction mechanisms associated with defects, barriers such as Schottky emission, Poole–Frenkel current, and hopping conduction are discussed in Section 4.4.4.

1.3 GaN Schottky Barriers for High-Voltage Rectifiers

The wide-bandgap nitride semiconductor family, particularly the GaN in combination with AlGaN system, is attractive from the viewpoint of fabricating high-voltage rectifier devices because of its large bandgap and relatively high electron mobility, the simplest example of which is the use of Schottky diodes [23–25]. The Schottky diodes have lower blocking voltages than p-i-n rectifiers, discussed in Section 4.7, but have advantages in terms of switching speed and lower forward voltage drop. To attain the bulk breakdown properties, edge termination techniques such as field rings on filed plates, bevels, or surface ion implantation are often employed. These were developed

for Si and SiC and maximize the high voltage blocking capability by avoiding sharp field distributions within the device. As the GaN development progresses, these techniques might be implemented in GaN Schottky diode rectifiers as well. Proper design of the edge termination is critical both for obtaining a high breakdown voltage and reducing the on-state voltage drop and switching time. Assuming a breakdown field of about 2×10^6 V cm^{-1} for GaN, one can sustain a voltage of 20 kV across 100 μm if the doping level is about 10^{15} cm^{-3} or less. The predicted breakdown field in GaN has been estimated at 2–3 × 10^6 V cm^{-1}.

Lateral Schottky barrier devices with large blocking voltages have been fabricated in GaN and have been grown on *c*-plane Al_2O_3 substrates by OMVPE using trimethylgallium and ammonia as the precursors. The reverse breakdown voltage of simple Schottky rectifiers fabricated on this material was >2 kV. A p-guard ring can be incorporated in the structure to reduce the field at the edges, which are otherwise higher than the interior of the device, as shown in Figure 1.21. The Schottky barrier metal is extended over an oxide layer at the edge to further minimize field crowding. The guard and field rings are formed by Mg^+ implantation followed by 1100 °C annealing. Without any edge termination, a breakdown voltage, V_B, is ~2300 V. The forward turn-on voltage is in the range 15–50 V, with a best on-resistance of 0.8 Ω cm^2. The figure-of-merit $(V_B)^2/R_{ON}$ is 6.8 MW cm^{-2}. As the guard-ring width

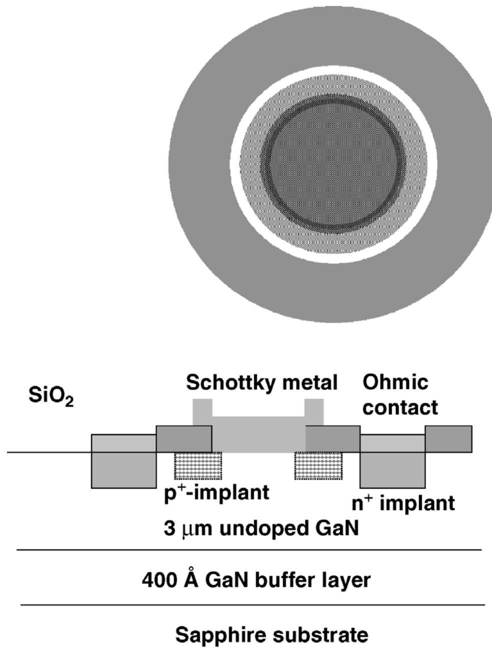


Figure 1.21 GaN power rectifiers with p-guard ring for edge terminations. Courtesy of S. J. Pearton [23–25].

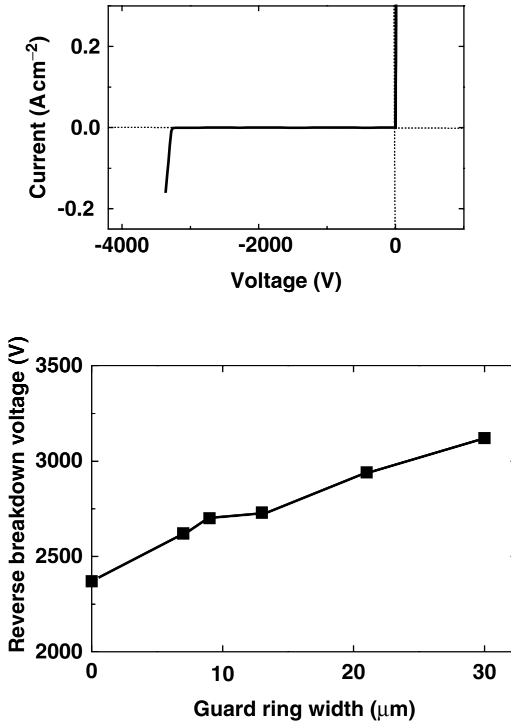


Figure 1.22 Current–voltage characteristics of GaN power rectifiers with p-guard ring for edge terminations (top) and the effect of p-guard ring width on the reverse breakdown voltage of GaN power rectifiers (bottom). Courtesy of S. J. Pearton [23–25].

was increased, a monotonic increase in V_B was observed, reaching a value of $\sim$3100 V for 30 μm wide rings, as shown in Figure 1.22. The figure-of-merit for this case is 15.5 MW cm^{-2} and the reverse leakage current is still in the nA range at voltages up to 90% of the breakdown value [23,24]. Additional spread of the field at the edges can be had by modifying the guard ring, which is floating, by incorporating additional guard rings, resulting in some improvement in V_B, albeit small as the main limitation is elsewhere.

If it could be grown with high enough quality and low enough unintentional doping, AlGaN with its increased bandgap should pave the way for even larger blocking voltages. Investigations of this very point with varying AlN mole fraction in the range of 0–0.25 have been carried out [23,24].

By doing so, a maximum V_B of 4.3 kV has been achieved for $Al_{0.25}Ga_{0.75}N$ diodes. At low reverse biases, these rectifiers typically show currents that are proportional to the contact perimeter, whereas at higher biases the current is proportional to contact area. The forward current characteristics show ideality factors of 2 at low bias (Shockley–Read–Hall recombination) and 1.5 at higher voltages (mixture of diffusion and generation–recombination).

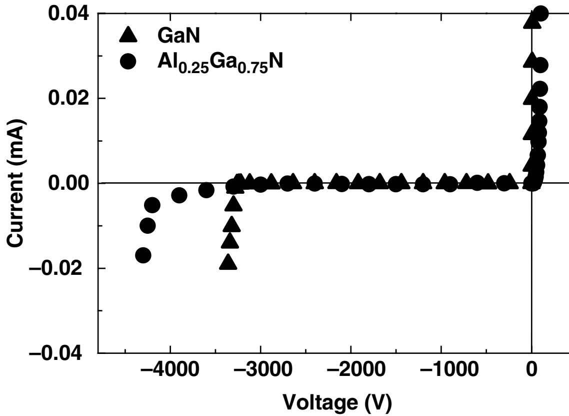


Figure 1.23 Room-temperature *I–V* characteristics from an $Al_{0.25}Ga_{0.75}N$ rectifier and for comparison from a GaN device. Courtesy of S. J. Pearton [23–25].

The undoped $Al_xGa_{1-x}N$ layers for this investigation were grown by atmospheric pressure organometallic vapor phase epitaxy, had an active layer thickness of $\sim$2.5 μm, and a resistivity on the order of $10^7\,\Omega\,cm$. In some cases, Si^+ was implanted at $5 \times 10^{14}\,cm^{-2}$, 50 keV, into the contact region and activated by annealing at 1150 °C for 10 s under N_2 to facilitate ohmic contacts. The contacts were then formed by liftoff of e-beam-evaporated Ti/Al/Pt/Au annealed at 700 °C for 30 s under N_2. The rectifying contacts were formed by liftoff of e-beam-evaporated Pt/Ti/Au (diameter 60–1100 μm). The devices were tested at room temperature under a Fluorinert ambient with the trend being that the breakdown voltage increased with increasing mole fraction up to the maximum mole fraction explored, which was 0.25. Figure 1.23 also shows an *I–V* characteristic from an $Al_{0.25}Ga_{0.75}N$ device lacking edge termination or surface passivation showing a V_B of about 4.3 kV. The on-resistance of the AlGaN diodes was higher than for pure GaN, owing to higher ohmic contract resistance. The lowest R_{ON} achieved was $3.2\,\Omega\,cm^2$, leading to a figure-of-merit of $\sim$5.5 MW cm^{-2}.

The breakdown voltage, V_B, can be affected by changing the gap between the Schottky barrier and the ohmic contact by employing p-guard rings and extending the Schottky contact edge over an oxide layer for edge termination. Doing so led to V_B values up to 9700 V for $Al_{0.25}Ga_{0.75}N$ rectifiers, with breakdown still occurring at the edges of the Schottky contact [23–25]. The reverse leakage current just before breakdown is dominated by bulk contributions in that it scales with the area of the rectifying contact. The material parameters and almost all the fabrication steps for these particular devices are the same as those for the above-mentioned rectifiers. The exception in fabrication is that with Mg^+ implantation at $5 \times 10^{14}\,cm^{-2}$, 50 keV, and 30 μm diameter p-guard rings, metal was formed at the edge of the Schottky barrier. The fabricated topology is very similar to that shown in Figure 1.21 with the exception of course of the material being AlGaN.

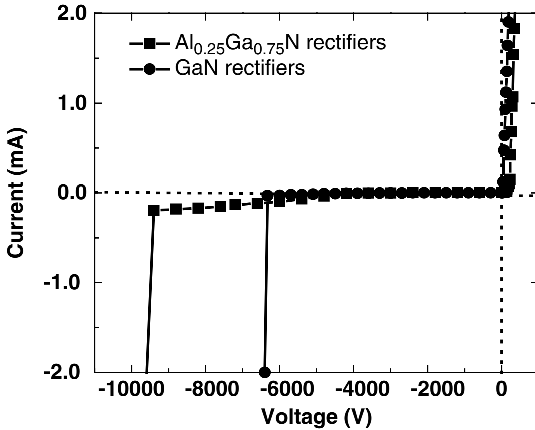


Figure 1.24 Comparison of the current–voltage characteristics for a GaN and an $Al_{0.25}Ga_{0.75}N$ rectifier Courtesy of S. J. Pearton [23–25].

Figure 1.24 shows *I–V* characteristics of GaN and $Al_{0.25}Ga_{0.75}N$ rectifiers with 100 μm gap spacing. The forward turn-on voltage, V_F, defined as that leading to a current density of $100\,A\,cm^{-2}$ is ~15 V for GaN and ~33 V for $Al_{0.25}Ga_{0.75}N$, indicating poor conductivity of the structures that is problematic in the on state. Figure 1.25 shows the measured V_B values for $Al_{0.25}Ga_{0.75}N$ and GaN (for comparison) rectifiers as a function of the gap spacing between the rectifying and ohmic contacts. For gaps between 40 and 100 μm, V_B varies nearly linearly with the spacing,

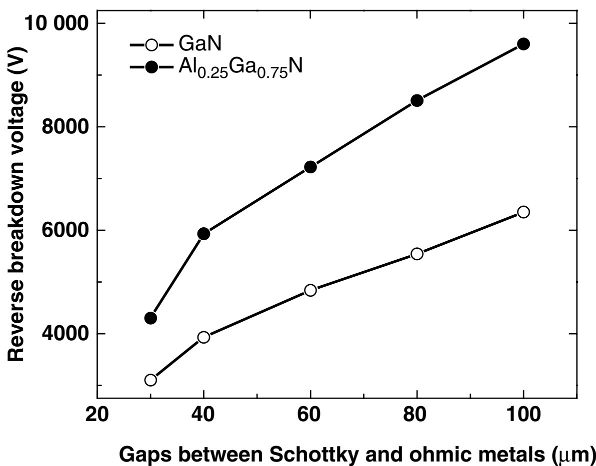


Figure 1.25 Effect of Schottky–ohmic contact gap spacing on V_B for GaN and $Al_{0.25}Ga_{0.75}N$ rectifiers. Courtesy of S. J. Pearton [23–25].

at slopes of $6.35 \times 10^5\,\mathrm{V\,cm^{-1}}$ for $Al_{0.25}Ga_{0.75}N$ and $4.0 \times 10^5\,\mathrm{V\,cm^{-1}}$ for GaN. The deviation observed at shorter spacing is owing to the fact that the 30 μm p-guard ring almost covers this region.

Because these types of devices are intended for high-power operation, which inevitably raises the junction temperature, there is a need to understand the current transport mechanisms, the origin of the reverse leakage current, and the magnitude and sign of the temperature coefficient for V_B. The investigations into this particular aspect of rectifiers indicated that over a broad range of voltages, the reverse leakage current is proportional to the diameter of the rectifying contact, indicating that surface periphery leakage is the dominant contributor [23–25]. The temperature coefficient for V_B was found to be negative for both GaN and AlGaN, even in edge-terminated devices, which is controversial as the temperature coefficient is positive in many well-established semiconductors.

1.4 Ohmic Contact Resistance

An ohmic contact is a metal–semiconductor contact that has a very small contact resistance compared to the bulk or spreading resistance of the semiconductor. It is said that the contact is ohmic when the ratio of the potential V drop across the contact versus the current I flowing through the contact is linear with a constant R_c. Ideal ohmic contacts should not contribute to the voltage drop across the device and should not alter the current–voltage relationship. Additionally, the contact must remain intact and robust regardless of the environment and the contact characteristics must not change with storage and dynamic operations. Naturally, not all of these requirements can be met simultaneously but gallant strides should be made to satisfy as many as possible.

1.4.1 Specific Contact Resistivity

Although the current–voltage (I–V) expression is sufficient, it is customary to deduce the specific resistance near-zero bias. Caution should be exercised, as the I–V characteristic may not be linear, thus causing a voltage-dependent resistance term. Nevertheless, the specific resistance creates an image of impediments to current flow. It is in this context that we define the specific contact resistivity, in terms of $\Omega\,\mathrm{cm}^2$. The product of R_c and the area A of the contact is called the specific contact resistance ρc expressed as

$$\rho_c = \left[\frac{\partial J}{\partial V}\right]^{-1}_{V=0} \quad (\Omega\,\mathrm{cm}^2). \tag{1.66}$$

For $kT/E_{00} \gg 1$ (*moderate doping concentrations*), the TE mechanism dominates the current conduction and the specific contact resistance near $V=0$ with the aid of

Equations 1.7 and 1.8 becomes

$$\rho_c = \frac{k}{qA^*T}\exp\left(\frac{q\phi_B}{kT}\right). \tag{1.67}$$

It clearly depends on temperature and at higher temperatures, there is more thermionic emission current, which results in a smaller ρc.

For $kT/E_{00} \approx 1$ (*intermediate doping concentrations*), a mixture of thermionic, thermionic field emission, and tunneling mechanisms is observed. The specific contact resistance can be obtained deriving Equation 1.10 with the use of Equation 1.11 with respect to voltage if thermionic emission is the dominating current conduction mechanism. Setting the voltage to zero and taking the Fermi energy at the conduction band edge leads to

$$\rho_c \approx \left(\frac{\partial J_{tef}}{\partial V}\right)^{-1}_{V=0} \approx \left\{\frac{A^*T^2q(\pi E_{00}q\phi_B)^{1/2}}{kT\cosh\frac{E_{00}}{kT}}\left[1-\frac{1}{2}\pi E_{00}(q\pi E_{00}\phi_B)\right]^{-1}\right\}^{-1} \times \exp\left(\frac{q\phi_B}{E_{00}\coth\frac{E_{00}}{kT}}\right), \tag{1.68}$$

or simply

$$\rho_c \propto e^{q\phi_B/(E_{00}coth(E_{00}/kT))}. \tag{1.69}$$

For $kT/E_{00} \ll 1$, which is associated with high doping levels, the tunneling current dominates and we have

$$\rho_c \propto \exp\left(\frac{q\phi_B}{E_{00}}\right). \tag{1.70}$$

In this case, ρc depends strongly on the doping concentration and the barrier. As the doping concentration is increased further, the depletion width of the Schottky junction decreases. This results in an increase of the tunneling transmission coefficient and a decrease of the resistance.

If a large number of surface states exist on the semiconductor surface, the barrier height is pinned at the semiconductor surface within its energy gap, and is independent of the metal work function. This is the Bardeen limit, which contrasts the Schottky limit where the metal–semiconductor contact is assumed ideal and the surface states are ignored [26,27]. In practice, the Fermi levels of most III–V compounds are pinned and the resultant barriers must be considered. The barrier height depends on the bandgap and the surface state density of the semiconductor.

1.4.2
Semiconductor Resistance

In addition to the metal–semiconductor resistance, the semiconductor resistance too must be added to the total resistance. The semiconductor resistance, owing mainly to

the neutral region, may be defined as [1]

$$R_s = \frac{1}{A_j} \int_{x_1}^{x_2} \rho(x) dx, \tag{1.71}$$

where x_1 represents the depletion edge, x_2 denotes the boundary of the epitaxial layer, $\rho(x)$ is the resistivity at x, and A_j is the area of the metal–semiconductor junction. The parameter x_1 depends on the depletion width W, which is a function of temperature,

$$W = \left[\frac{2\varepsilon_s}{qN_{\text{Deff}}} \left(V_{\text{bi}} - V - \frac{kT}{q} \right) \right]^{1/2}, \tag{1.72}$$

where V_{bi} is the built-in potential given by

$$V_{\text{bi}} = \phi_B - \eta = \phi_B - \frac{kT}{q} \ln\left(\frac{N_c}{N_D} \right). \tag{1.73}$$

In GaN-based p–n junctions, such as those in LEDs and lasers, p-type contact resistance dominates because of a large metal–semiconductor barrier and a large effective mass. In addition, the semiconductor resistance is also large because of a combination of the low hole concentration and mobility. The nonohmic behavior caused by a combination of a high metal-to-p-semiconductor barrier and a low hole concentration, unless chemical interaction between the metal layers and the semiconductor causes the direct tunneling-like current to dominate, will give rise to a voltage drop as well as an increased resistance exacerbating the Joule heating and the resultant rise in junction temperature. In addition, sapphire substrates, must they be used, are semi-insulating, necessitating the use of surface contacts for both n- and p-type regions. Doing so requires current conduction laterally from the n-type contact to the junction area. Owing to the considerable distance involved between the metal contact and the junction area, the semiconductor resistance is considerable. It could be lowered if the n-type semiconductor is sufficiently thick. What is exasperating is that highly Si-doped n-type GaN cracks if its thickness is increased beyond about 3 μm because of the residual thermal strain.

1.5 Determination of the Contact Resistivity

The most widely used method for determining the *specific contact resistance* is the method of *transfer length* first introduced by Shockley [28]. Called transmission line model (TLM), this method was refined and/or expanded later on by Murrmann and Widmann [29], Reeves and Harbison [30], and Berger [31,32]. See Ref. [33] for a full treatment. In this particular approach, a linear array of contacts is fabricated

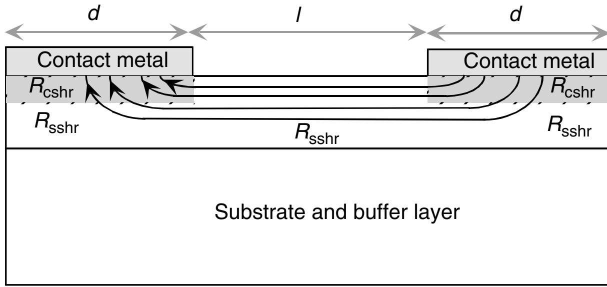


Figure 1.26 Schematic representation of alloyed ohmic contacts to a semiconductor where the filled region below the metallization indicates the altered semiconductor owing to an interaction between the contact material and the semiconductor. Here R_{sshr} and R_{cshr} indicate the sheet resistance of the unaltered semiconductor and the sheet resistance of the altered semiconductor below the contact metal after annealing.

with various spacing between them. The pattern used and the resistance versus the gap spacing l (l_{12}, l_{23}, l_{34}, etc.) are depicted in Figure 1.26. The total resistance is given by

$$R_T = 2R_C + \frac{lR_{sshr}}{Z}, \tag{1.74}$$

where the first term represents twice the contact resistance R_c because the resistance is measured between two identical contacts, whereas the second term is because of the semiconductor resistance that depends on the contact separation or the gap between contacts. The term R_{sshr} denotes the sheet resistance of the semiconductor layer. Care must be taken to account for the resistance between the ohmic contacts and measurement setup. If an overlay metallization is used and the probe to the overlay metallization contact is very good, then resistance can be negligible. The contact resistivity measurement methods that have been devised not only deduce that resistivity but also the semiconductor sheet resistance.

Contacts formed on heavily doped semiconductor that take advantage of tunneling do not alter the semiconductor properties under the metal. However, in contacts relying on interaction of the deposited metal and the underlying semiconductor through annealing, properties of the semiconductor and thus its resistance under the contact metal are altered. Consequently, the semiconductor resistance under the contact differs from that of the outside the contact region as shown in Figure 1.26, where R_{sshr} and R_{cshr} represent the semiconductor sheet resistance and sheet resistance of the alloyed region under the contact metal, respectively. In addition, the current flow paths are also shown. Because the current flows through the least path of resistance, the current density is higher near the edge of the contact.

Referring to Figure 1.26, the current distribution in the alloyed region can be expressed as

$$dI(x) = -ZJ(x)dx, \tag{1.75}$$

where Z is the width of the ohmic contact, x represents the distance parallel to the surface of the semiconductor, and J is the current density, which can be related to the specific contact resistance through [34]

$$J(x) = \frac{V(x)}{r_\mathrm{c}}, \tag{1.76}$$

where r_c is the specific contact resistivity and $V(x)$ is the channel potential with respect to the potential of the contact metal. The voltage distribution is expressed as

$$\mathrm{d}V(x) = -\frac{I(x)R_\mathrm{cshr}}{Z}\mathrm{d}x, \tag{1.77}$$

where R_cshr is the sheet resistance per square of the region under the ohmic contact.

Equations 1.75 and 1.77 may be reduced to that describing a transmission line model in electromagnetics as

$$\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} = \frac{V}{(L_\mathrm{T})^2}, \tag{1.78}$$

where L_T represents the transfer length defined as

$$(L_\mathrm{T})^2 = \frac{r_\mathrm{c}}{R_\mathrm{cshr}}. \tag{1.79}$$

Equation 1.78 holds if the epitaxial layer thickness is much smaller than L_T. For solving Equation 1.78 the boundary conditions are

$$\left.\frac{\mathrm{d}V}{\mathrm{d}x}\right|_{x=0} = I_0\frac{R_\mathrm{cshr}}{Z} \quad \text{and} \quad \left.\frac{\mathrm{d}V}{\mathrm{d}x}\right|_{x=d} = 0, \tag{1.80}$$

where d is the contact length. The second boundary condition states that the electric field in the alloyed region at the end of the contact opposite of the channel region is zero. The second order differential equation represented by Equation 1.78 has a general solution of the form

$$V(x) = C_1\exp\left(\frac{x}{L_\mathrm{T}}\right) + C_2\exp\left(\frac{-x}{L_T}\right). \tag{1.81}$$

And application of the boundary conditions of Equation 1.80 results in the determination of constants C_1 and C_2 as

$$C_1 = \frac{I_0 R_\mathrm{cshr} L_\mathrm{T}\exp(d/L_\mathrm{T})}{Z[\exp(d/L_T) - \exp(-d/L_\mathrm{T})]},$$
$$C_2 = \frac{I_0 R_\mathrm{cshr} L_T\exp(-d/L_\mathrm{T})}{Z[\exp(d/L_\mathrm{T}) - \exp(-d/L_T)]}. \tag{1.82}$$

Solving for V in Equation 1.81 at $x = 0$ results in

$$V(0) = I_0 R_\mathrm{cshr}\left(\frac{d}{Z}\right)\left(\frac{L_\mathrm{T}}{d}\right)\coth\left(\frac{d}{L_T}\right), \tag{1.83}$$

which can be conveniently expressed as

$$V(0) = I_0 R_c, \tag{1.84}$$

where R_c is the contact resistance and is related to transfer length and sheet resistance in the contact region through $R_c = R_{cshr}(d/Z)F_{tlm}$ and $F_{tlm} = (L_T/d)\coth(d/L_T)$.

For $d/L_T \ll 1$, $\coth(d/L_T) \approx L_T/d$, which leads to $F_{tlm} \approx (L_T/d)^2$ and

$$R_c = R_{cshr}\left(\frac{L_T^2}{dZ}\right). \tag{1.85}$$

For $d/L_T \gg 1$, which is the case for many of the patterns employed, $F_{tlm} \approx L_T/d$ and

$$R_c = R_{cshr}\left(\frac{L_T}{Z}\right). \tag{1.86}$$

Solving for V in Equation 1.81 at $x = d$ results in

$$V(d) = \frac{I_0 R_{cshr} L_T}{Z\sinh(d/L_T)}. \tag{1.87}$$

An end resistance R_{end} can be defined as

$$R_{end} = \frac{V(d)}{I_0}, \tag{1.88}$$

and calculated using Equation 1.87 as

$$R_{end} = \frac{R_{cshr} L_T}{Z\sinh(d/L_T)}, \tag{1.89}$$

using $R_c = R_{cshr}(L_T/Z)\coth(d/L_T)$, and

$$\frac{R_c}{R_{end}} = \cosh\left(\frac{d}{L_T}\right). \tag{1.90}$$

The resistivity under the contact region can be determined. Equations 1.88 and 1.90 are used in the measurements and calculation of ohmic contact resistivity.

Exercise

A useful exercise would actually be to go through an example of how the contact resistivity is found. In this vein, let us calculate the values for the specific contact resistivity, the transfer length, semiconductor sheet resistance, and the semiconductor sheet resistance underneath the ohmic contacts with the end resistance $R_{end} = 0.02\,\Omega$, $Z = 100\,\mu\text{m}$, and $d = 50\,\mu\text{m}$. Let us also assume that we actually performed resistance measurements versus gap in a TLM pattern and found the contact resistance to be $R_c = 0.05\,\Omega$. Beginning with Equation 1.89 we determine the sheet resistance underneath the ohmic contact region paving the way to the transfer length, which then leads to the contact resistivity as detailed in the step-by-step

treatment given below:

$$R_{\mathrm{end}} = \frac{R_{\mathrm{cshr}} L_{\mathrm{T}}}{Z \sinh(d/L_{\mathrm{T}})}$$

and thus

$$R_{\mathrm{cshr}} = \frac{R_{\mathrm{end}} Z \sinh(d/L_{\mathrm{T}})}{L_{\mathrm{T}}};$$

knowing R_{end} one gets

$$R_{\mathrm{cshr}} = 0.14\,\Omega,$$

$$L_{\mathrm{T}} = \frac{d}{\cosh^{-1}(R_{\mathrm{c}}/R_{\mathrm{end}})} = 31.9\,\mu\mathrm{m},$$

$$r_{\mathrm{c}} = L_{\mathrm{T}}^2 R_{\mathrm{cshr}} = 1.46 \times 10^{-6}\,\Omega\,\mathrm{cm}^2,$$

or in terms of the measurable quantities

$$r_{\mathrm{c}} = \frac{d^2}{[\cosh^{-1}(R_c/R_{\mathrm{end}})]^2} \frac{R_c Z}{L_{\mathrm{T}} \coth(d/L_T)} = 1.46 \times 10^{-6}\,(\Omega\,\mathrm{cm}^2).$$

In the TLM method, the intercontact spacing (gap) must be much smaller compared to contact width to avoid edge effects. If not, the electric field near the edges would not be fully along the gap and would have a normal component as well. This problem is completely eliminated by using the circular transmission line method (CTLM) where the contacts are circular, as opposed to rectangular, and concentric. In this case, the total resistance is approximately given by, assuming the metal sheet resistance to be negligible [35–37]:

$$R = \left(\frac{R_{\mathrm{sshr}}}{2\pi}\right)\left[\ln\left(\frac{r_0}{r_1}\right) + L_{\mathrm{T}}\left(\frac{1}{r_0} + \frac{1}{r_1}\right)\right], \tag{1.91}$$

where r_{i} is the inner contact pad radius and r_0 is the outer contact pad radius that is r_{i} plus the relative spacing. As in the case of standard TLM method, extrapolation of the spacing to zero, meaning setting $r_{\mathrm{i}} = r_0$, leads the resistance to be equal to twice the contact resistance from which one can determine the specific contact resistivity. Here the implicit assumption is that the sheet resistance under the contact metal is the same as the unperturbed semiconductor sheet resistance. Even with these improved geometries, the TLM method is not accurate for specific contact resistivities near and below $10^{-7}\,\Omega\,\mathrm{cm}^2$, in which case the Kelvin probe measurements can be employed.

In the four-terminal (Kelvin) resistor method [38], the test pattern consists of four metal pads on an insulator. Two are connected to a semiconductor bar by means of large-area contacts. The other two touch the semiconductor at the contact opening. One pair is used to pass current whereas the other two are used to measure the voltage drop. In contrast to the TLM case, the resistance of the line outside the contact area does not contribute to the contact resistance in this test structure; this allows a more precise measurement. For very small contact resistances, this method is more accurate than the TLM method.

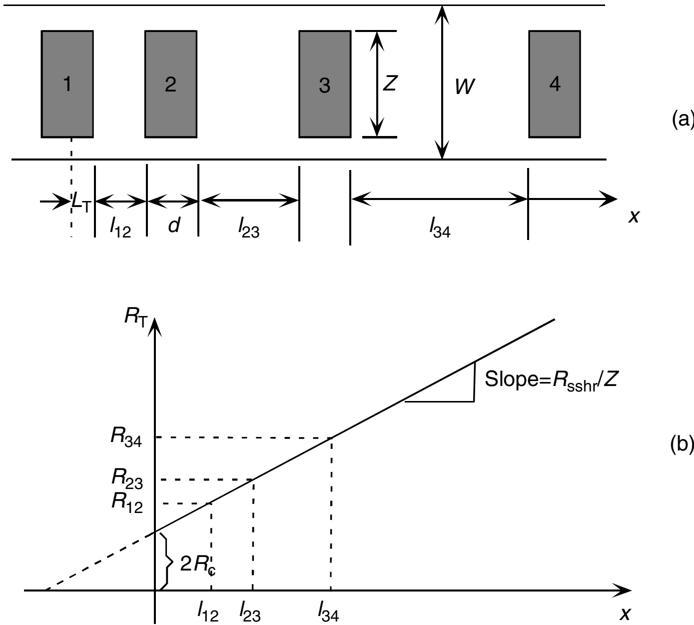


Figure 1.27 (a) Top view of a transmission line pattern commonly used to deduce the specific contact resistivity in planar contacts. (b) The variation of the resistance with respect to the gap distance.

The value of R_c can be determined from the intercept of the resistance, R, versus l measurements. The semiconductor sheet resistance, R_{sshr}, can be found from the slope of R versus l line. When the sheet resistance of the active layer below the annealed region is different from that of the semiconductor in between the contacts and additional measurement is needed to determine sheet resistance in the alloy region and r_c. Even though this is not likely the case for n-type contacts in nitride semiconductors as they are surface oriented, the same may not be true for p-type contacts considering the very defective nature of p-type GaN and expected penetration of the ohmic contact metal. The end resistance can be determined by measuring the current–voltage characteristics of the sample structure shown in Figure 1.27 using [34]

$$R_{end} = \frac{V_{2,3}}{I_{1,2}}, \tag{1.92}$$

where $V_{2,3}$ is the voltage between contacts 2 and 3 and $I_{1,2}$ is the current flowing between contacts 1 and 2. Contact and end resistances, R_c and R_{ernd}, can be found using Equation 1.90. From Equation 1.82 one can express the transfer length in terms of those two resistances as

$$L_T = \frac{d}{\cosh^{-1}(R_c/R_{end})}. \tag{1.93}$$

R_{cshr} and r_c can be found using Equation 1.79 together with Equation 1.90.

For contact resistivities in other semiconductors and pertinent technological issues, the curious reader is referred to Scorzoni and Finetti [39] and Shen *et al.* [40] who treated the contact resistivity and its measurement in detail. In addition to defining the relation between contact resistance and contact resistivity, the researchers thoroughly described different types of approaches for improving the measurement techniques. They also noted the mixed use of the terminology contact resistivity, specific contact resistivity, specific contact resistance, specific interface resistance, specific resistance, and contact resistance for the same term. Here, the term specific contact resistivity has been adopted.

1.6 Ohmic Contacts to GaN

Early studies suggested that although an ohmic contact is possible for n-GaN using Al and Au metallization, the specific contact resistivity is relatively high, about 10^{-4}–10^{-3} cm^2 [41]. To form an ohmic contact, metals were deposited by evaporation and then patterned by photolithography and liftoff techniques. From current–voltage measurements it was determined that the *as-deposited* Al contacts were ohmic. However, the as-deposited Au contacts were rectifying and became ohmic after annealing at 575 °C, paving the way for the argument that work functions of GaN and Al are very close to one another and that the surface of GaN is defect free. Using Ti/Al to contact n-type GaN, the specific contact resistivity was lowered to about $8 \times 10^{-6}\,\Omega\,cm^2$ by annealing at 900 °C for 30 s [42]. Au (and later Au/Ni) and Ti/Al were utilized as p-type and n-type contacts, respectively, in LED structures [43,44]. Although no contact resistance was reported, the LED operating voltages of 4 V at 20 mA forward bias clearly demonstrates that the contact resistance was reasonably low.

1.6.1 Nonalloyed Ohmic Contacts

Lin *et al.* [45] developed a novel scheme of nonalloyed ohmic contacts on GaN employing a 10 period InN/GaN (10 nm/l0 nm) short-period superlattice (SPS), as shown in Figure 1.28. The GaN and InN films were doped to about $5 \times 10^{18}\,cm^{-3}$. The background electron concentration in InN film may be higher though. TLM measurements revealed a specific contact resistivity as low as $6 \times 10^{-5}\,\Omega\,cm^2$. The current as a function of the applied gate voltage (Figure 1.29) revealed strong rectifying features with only an InN cap layer but good ohmic characteristics with the short-period superlattice structures. A speculation suggested that a low Schottky barrier contact on InN together with an increasing tunneling transmission coefficient may be responsible for the low-resistance contacts, as it was demonstrated for InGaAs [46]. The tunneling parameter E_{00} is about 0.37 eV for $N_D = 5 \times 10^{18}\,cm^{-3}$. Because E_{00} is much larger than the thermal energy kT at room temperature, the possibility of tunneling as the dominant mechanism underlying electron transport through a GaN/InN superlattice ohmic contact is very real. InGaN/GaN supperlattices have also been utilized to reduce the resistance of metal contact to p-GaN.

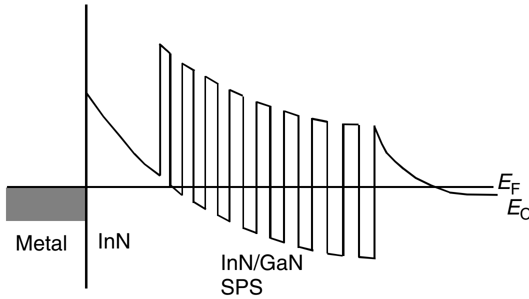


Figure 1.28 Schematic band structure of an InN/GaN short-period superlattice employed to achieve contacts on n-type GaN.

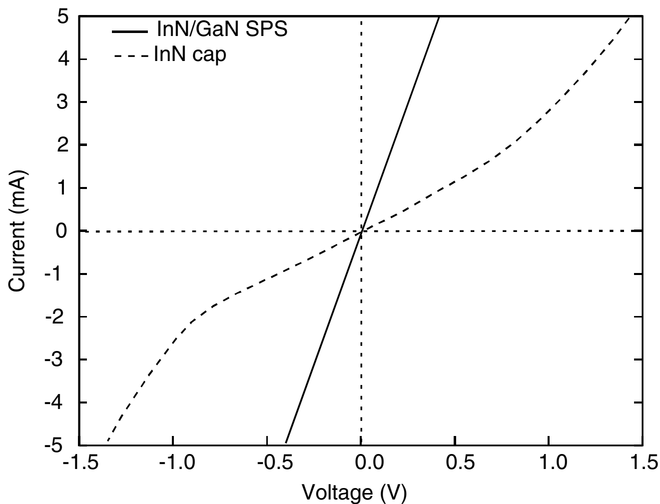


Figure 1.29 Current–voltage characteristics of a metal–GaN system with an InN surface layer (dashed line) and of that with InN/GaN SPS surface layer (solid line).

Kumakura *et al.* [47] demonstrated that a 2 nm thick $In_{0.19}Ga_{0.81}N$ contact layer (Mg doped, 4×10^{19}) inserted between Pd/Au and p-GaN can achieve a low contact resistance of $9.3 \times 10^{-6}\,\Omega\,cm^2$ without any special treatment. Specific contact resistance of $6 \times 10^{-3}\,\Omega\,cm^2$ was achieved for Ni/Au contact on thin $In_{0.27}Ga_{0.73}N$ to p-GaN without any optimization [48]. The polarization field and electric field due to the ionized acceptors in the surface depletion layer decrease the tunneling barrier width and enhance the tunneling transport, resulting in the reduction of contact resistance.

1.6.2
Alloyed Ohmic Contacts on n-Type GaN

During the past few years, several attempts have been made to obtain low-resistance ohmic contacts on GaN [41,42,45,49]. Initial attempts of Foresi and Moustakas [41],

who used Au or Al after annealing at 575 °C, led to contacts with a resistivity of $\rho_c \approx 10^{-6}\,\Omega\,\text{cm}^2$. Later, a specific contact resistance of $1 \times 10^{-4}\,\Omega\,\text{cm}^2$ was obtained for tungsten ohmic contacts to highly doped n-GaN [50]. Contacts on n-GaN were later improved significantly by Lin *et al.* [42]. They employed a Ti/Al bilayer deposited via conventional electron-beam evaporation on a GaN epitaxial layer followed by thermal annealing at 900 °C for 30 s in an N_2 ambient (by rapid thermal annealing (RTA)). This Ti/Al metallization yielded $\rho_c = 8 \times 10^{-6}\,\Omega\,\text{cm}^2$ but suffered from Ga outdiffusion and the subsequent reaction with Al rendering the surface metal discontinuous and have a high resistivity. To minimize the high-resistivity problem, Wu *et al.* [49] added a second set of Ti/Al stack, following the annealing step, which requires realignment in conjunction with a second photolithography step. The ohmic contact resistivity was lowered to $\rho_c \approx 3 \times 10^{-6}\,\Omega\,\text{cm}^2$. Despite the oxidation and the resultant high metal resistivity, the Ti/Al metallization schemes have actually been investigated reasonably extensively [51–54].

Building on the earlier work of Lin *et al.* [42], a multilayer ohmic contact to n-GaN based on Ti/Al metallization has been designed by Fan *et al.* [55]. The GaN film utilized were grown by molecular beam epitaxy (MBE) and doped with Si to a level of $(1\text{–}5) \times 10^{17}\,\text{cm}^{-3}$. The GaN surface was etched first with Cl_2 for 20 s, a BCl_3 etched for another 20 s, and thereafter a composite metal layer of Ti/Al/Ni/Au (150/2200/400/500 Å) was deposited. It should be mentioned that Ni serves as barrier also keeping Al and Au from reacting with each other. Two measurements of the specific resistivity were carried out, before and after RTA treatments. Measurements of the electrical current through the metal contacts to n-GaN as a function of voltage were performed before (*solid line*) and after (*dashed line*) alloying (Figure 1.30). These

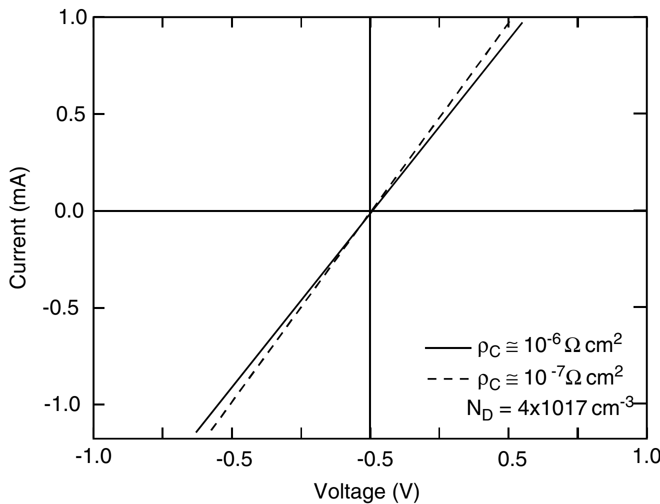


Figure 1.30 Current–voltage characteristics for nonalloyed (solid line) and alloyed (dashed line) metal–n-GaN contacts. The GaN surface was subjected to a <reactive ion etching, which is believed to cause a damage-induced increase of the electron concentration on the surface.

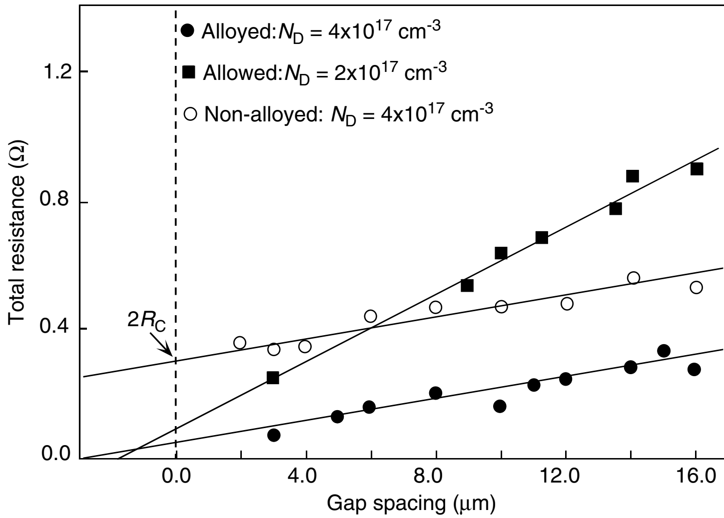


Figure 1.31 Total resistance versus the gap spacing for alloyed and nonalloyed contacts on GaN shown in Figure 1.30.

measurements indicate that the variation of the current with voltage is nearly linear even at sufficiently large current levels (100 mA) both in alloyed and nonalloyed ohmic contacts. Only annealing reduced the barrier height slightly after RTA, thus resulting in a small increase in current.

Contact resistances were derived from the plot of measured resistance versus gap spacing by TLM. The total resistance R_T between the two contacts was determined at room temperature employing a four-point probe arrangement (the contact resistivity ρ_c was derived from a plot of R_T versus gap length in Equation 1.74). The method of least squares was utilized to fit a straight line to the experimental data (Figure 1.31). The specific contact resistivity depends both on the annealing and the doping concentration of the semiconductor sample. The contact resistivity ρc was about $(9\text{–}12) \times 10^{-8}\,\Omega\,\text{cm}^2$ for doping levels between 4×10^{17} and $2 \times 10^{17}\,\text{cm}^{-3}$. For a doping level of $4 \times 10^{17}\,\text{cm}^{-3}$, alloying reduced the specific contact resistivity by about a factor of 40.

The determination of the contact resistivity was based on the assumption that the semiconductor sheet resistance underneath the contacts remains unchanged, Equation 1.79 with the modification that the sheet resistance under the contact region R_{cshr} is equal to the sheet resistance of the semiconductor between the contact regions R_{cshr} that is not true for alloyed contacts (see Figure 1.26). As mentioned previously, a metal with a low enough work function does not exist for good ohmic contacts on n-type GaN. The unavoidably large barriers diminish the possibility of thermionic emission governed ohmic contacts to GaN. The alternative mechanism is naturally some form of tunneling that may take place if GaN is so heavily doped to cause a very thin depletion region. Tunneling is possible if, owing to annealing, for example, at 900 °C for 30 s, Al and Ti along with Ni undergo substantial interaction with each

other and GaN. A cursory look would imply that Ti receives N from GaN forming a metallic layer whereas the lack of N in GaN provides the desired benefit of an increased electron concentration through N vacancy formation. Al acts to passivate the surface and also possibly reacts with Ti forming TiAl. This very possibility was investigated in contacts formed by TiAl and TiAlNiAu before and after annealing by Rutherford backscattering (RBS) by Duxstad *et al.* [56]. Annealed and unannealed Ti/Al contacts on GaN, examined by RBS, indicate some Ti and GaN on the surface (Figure 1.32), which leads to the conclusion that the picture is more complicated. Similar experiments with the multilayered Ti/Al/Ni/Au (500/2000/700/600 Å) contacts revealed the presence of Ti, Al, and Ni on the surface after annealing (Figure 1.33).

The oxygen present in the annealing ambient has a detrimental effect on the resulting electrical properties of Al/Ti/n-GaN contacts [54]. A Ni/Au bilayer directly on the top of Ti/Al addresses the oxidation issue with much success, as has been demonstrated in Ref. [55]. Diffusion barriers such as Pt in the form of Ti/Al/Pt/Au contacts [57–60] have been investigated with a contact resistivity of as low as 0.039 Ω mm corresponding to a specific contact resistance of about $5 \times 10^{-8}\,\Omega\,\text{cm}^2$ [61]. The precleaning procedure included ion etching in O_2 plasma followed by a rinsing step in dilute NH_4OH, which is known to etch GaN, particularly damaged GaN very slightly. It should be pointed out that the plasma damage prior to deposition is a concept developed by Fan *et al.* [55] and the TLM patterns are inaccurate for low ohmic contact resistivities. However, the 0.039 Ω mm is a good representative of the contact resistance. However, later studies indicated Pt not to be so effective a diffusion barrier. In fact, when Pt with its larger work function as compared to Au (which is one reason for the higher Schottky barrier heights obtained with Pt on n-type GaN) diffuses, it exacerbates ohmic contacts if subjected to high tempera-

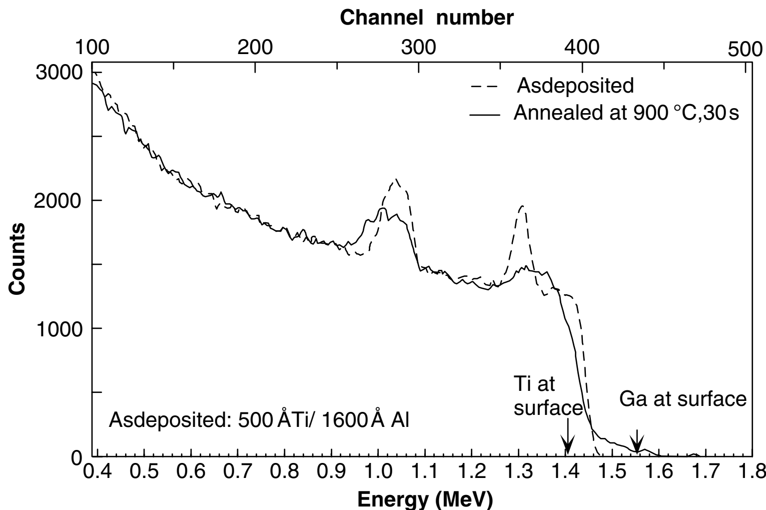


Figure 1.32 Rutherford backscattering profile of TiAl contacts on GaN before and after annealing [56].

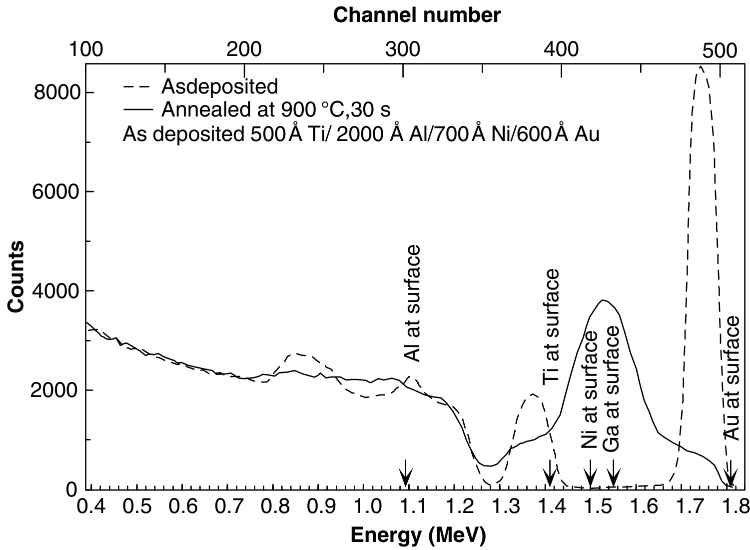


Figure 1.33 RBS profile of Ti/Al/Ni/Au contacts on GaN before and after annealing [56].

tures [62]. After a 450 °C-anneal, Pt was shown to diffuse into the GaN accompanied with Ga outdiffusion from GaN substrate as determined by RBS measurements. Moreover, Ga_3Pt_5 have been detected in the synchrotron XRD scans for samples annealed at temperature equal to or higher than 450 °C, which has been cited as the cause for ohmic contact degradation. In the investigation of Shen *at al.* [60], probing the interface between the contact and the AlGaN on which the contact was deposited AlTi, $AlTi_2$, and $AlTi_3$ were found for samples annealed at 700 °C by X-ray analyses. Moreover, Ti_2N and Au_3Pt were detected following annealing at 800 and 900 °C, respectively.

Ti/Al/Pd/Au [63], Ti/Al/Ti/Au [64], Ti/Al/Mo/Au [65], Ti/Al/Nb/Au [66], Ti/Al/$TiAl_3$ [67], and Ta/Ti/Al/Ni/Au [68] have also been examined for either better resistivities, thermal stability, better postanneal surface morphology and/or oxidation resistance layer. Ti/Al/Pt/Au (200/800/400/1500 Å) annealed at 900 °C for 35 s in N_2 gas have been reported to yield [69]. As in the case of Ti/Al/Ni/Au contacts annealed at nearly 900 °C, Ta/Ti/Al/Ti/Au contacts, requiring about 700 °C annealing temperature, led to ohmic contact resistivities of about $7.5 \times 10^{-7}\,\Omega\,cm^2$ on AlGaN/GaN heterostructures [68]. Downward or upward deviation in annealing temperature rapidly increased the ohmic contact resistance. The nitride phases have also been identified as the TaN/TiN for Au/Ni/Al/Ti/Ta and TiN for the Au/Ni/Al/Ti contacts. It should be noted that annealing of Ti/Al/$TiAl_3$ can also be carried out at lower temperatures (700 °C) and the contact resistance is very insensitive to the amount of oxygen in the annealing ambient. This metallization scheme has been touted a potential candidate for implementation in the devices that require both n-GaN and p-GaN contacts on the same wafer, because it can achieve near optimal performance

in an oxidation anneal at roughly the same temperature as is common for Ni/Au contacts to p-GaN annealed in an oxidizing atmosphere [67]. Other techniques such as vacuum annealing [70], minimizing the oxidation of the Al surface, and Cr/Al/Ni/Au requiring considerably lower annealing temperature (550–600 °C) [71] have also been explored [72]. The achieved specific contact resistances are about $1–3 \times 10^{-5}\,\Omega\,cm^2$.

The morphology of ohmic contacts is also an important parameter as it signals phase changes and also affects the line edge definition for the subsequent lithographic processes. Ti/Al/Ti/Au ohmic contacts result in relatively poor surface morphology following annealing. However, as in the case of Ti/Al/Ni/Au contacts annealed at nearly 900 °C, Ti/Al/Ti/Au contacts, requiring about 700 °C annealing temperature, led to ohmic contact resistivities of about $7.5 \times 10^{-7}\,\Omega\,cm^2$ on AlGaN/GaN heterostructures [68]. The nitride phases have also been identified as the TaN/TiN for Au/Ni/Al/Ti/Ta and TiN for the Au/Ni/Al/Ti contacts. The main drawback of this contact system as compared to the Ti/Al/Ni/Au system is that the lateral flow of the melt during annealing causes line definition degradation reaching a severe case when annealed, for example, at 850 °C for 30 s. If, on the contrary, the Ti barrier layer between Al and Au is replaced with a refractory metal such as Mo, the Al–Au alloy formation that is endemic in TiAlTiAu system is suppressed in the Ti/Al/Mo/Au system. With 850 °C annealing, the Mo barrier system leads to contact resistances as low as approximately $0.15\,\Omega\,cm$, with the specific contact resistance dipping down to the low $10^{-6}\,\Omega\,cm^2$ range. The important feature of the Mo barrier is that the resulting surface morphology of the contacts possesses good line definition.

It is difficult to achieve a low contact resistance on AlGaN owing to the large bandgap of AlN and hence large Schottky barrier height of many metals on AlGaN. By thinning the AlGaN barrier in AlGaN/GaN heterojunction field effect transistors by selective growth or selective etching, an ohmic contact of low resistivity can be obtained because of the carrier-tunneling enhancement in these regions. However, it requires multiple processing steps. By using "advancing interface reaction," the resistance of Ti/Al contacts on AlGaN/GaN structures can be reduced to $5.6 \times 10^{-6}\,\Omega\,cm^2$, where nitride forming metal Ti is used with the help of direct Si implantation [73]. It was expected that the thickness of Ti should increase with the increase of AlGaN thickness to obtain low resistance. Part of Ti reacts with Al to form Ti_3Al, whereas the remaining Ti reacts with AlGaN to form $AlTi_2N$. The depletion of AlGaN, which hinders carrier tunneling, leads to the reduced specific contact resistivity. However, there are two drawbacks of Ti/Al-based metallization for AlGaN/GaN MODFET: the requirement of a capping layer for annealing and low backscattered electron emission owing to low atomic number of Al and Ti. To overcome these drawbacks, Qiao *et al.* [74] proposed a Ta-based interface ohmic contact, which leads to a resistivity of $5 \times 10^{-7}\,\Omega\,cm^2$, efficient electron emission for e-beam lithographic alignment, and elimination of the capping layer for the ohmic annealing. The specific contact resistances of Ti/Al/Ni/Au ohmic contacts to AlGaN/GaN MODFET structures, with the value of $3.44 \times 10^{-6}\,\Omega\,cm^2$ [75] and 7.3×10^{-7} cm^2 [76], have also been reported.