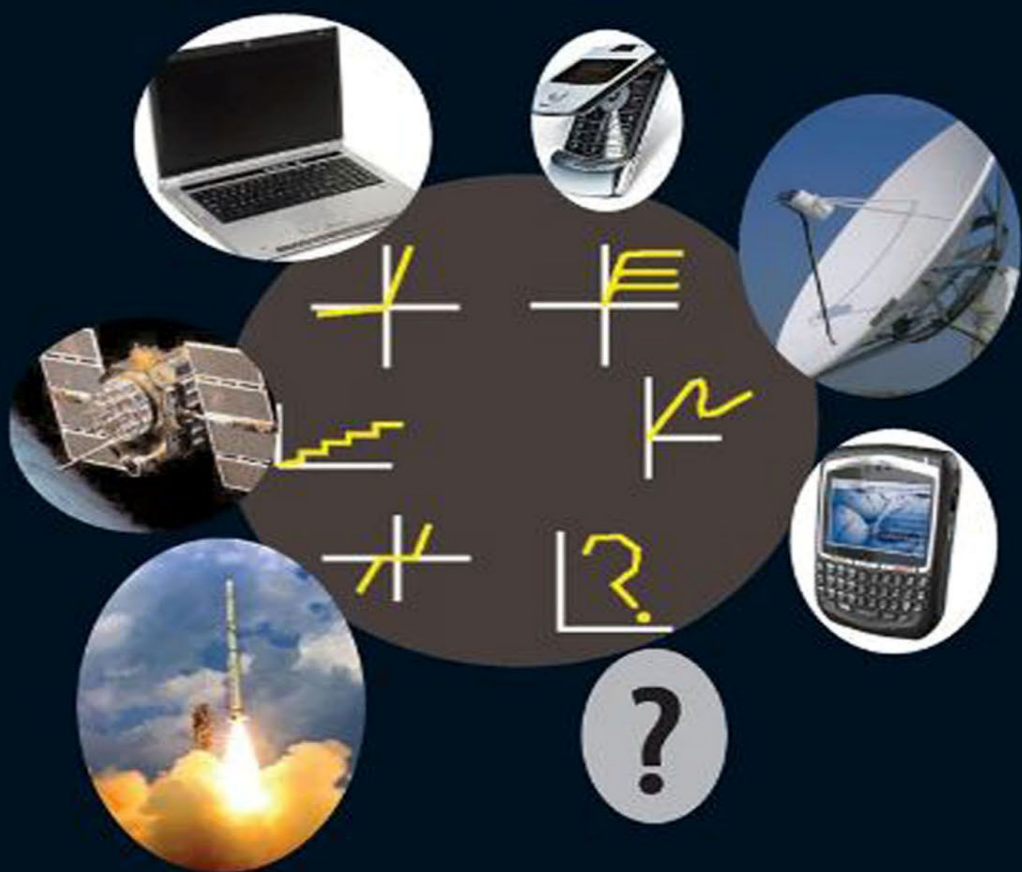


# Semiconductor Device Physics and Design



Umesh K. Mishra  
Jasprit Singh

 Springer

# Semiconductor Device Physics and Design

*by*

UMESH K. MISHRA

*University of California, Santa Barbara, CA, USA*

*and*

JASPRIT SINGH

*The University of Michigan, Ann Arbor, MI, USA*



Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-6480-7 (HB)  
ISBN 978-1-4020-6481-4 (e-book)

---

Published by Springer,  
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

*www.springer.com*

*Printed on acid-free paper*

All Rights Reserved

© 2008 Springer

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

*This book is dedicated to our parents  
Srinibas Mishra and Sushila Devi  
Gurcharn Singh and Gursharan Kaur*

# CONTENTS

|                                                                         |             |
|-------------------------------------------------------------------------|-------------|
| <b>Acknowledgements</b>                                                 | <b>xii</b>  |
| <b>Preface</b>                                                          | <b>xiii</b> |
| <b>Introduction</b>                                                     | <b>xx</b>   |
| <b>1 Structural Properties of Semiconductors</b>                        | <b>1</b>    |
| 1.1 INTRODUCTION                                                        | 1           |
| 1.2 CRYSTAL STRUCTURE                                                   | 2           |
| 1.2.1 Basic Lattice Types                                               | 2           |
| 1.2.2 Basic Crystal Structures                                          | 3           |
| 1.2.3 Notation to Denote Planes and Points in a Lattice: Miller Indices | 7           |
| 1.2.4 Artificial Structures: Superlattices and Quantum Wells            | 11          |
| 1.2.5 Surfaces : Ideal Versus Real                                      | 12          |
| 1.2.6 Interfaces                                                        | 15          |
| 1.2.7 Semiconductor Defects                                             | 15          |
| 1.3 LATTICE MISMATCHED STRUCTURES                                       | 17          |
| 1.4 STRAINED EPITAXY: STRAIN TENSOR                                     | 21          |
| 1.5 TECHNOLOGY CHALLENGES                                               | 24          |
| 1.6 PROBLEMS                                                            | 25          |
| 1.7 FURTHER READING                                                     | 27          |
| <b>2 Electronic levels in semiconductors</b>                            | <b>28</b>   |
| 2.1 INTRODUCTION                                                        | 28          |
| 2.2 PARTICLES IN AN ATTRACTIVE POTENTIAL:<br>BOUND STATES               | 29          |
| 2.2.1 Electronic levels in a hydrogen atom                              | 29          |
| 2.2.2 Electrons in a quantum well                                       | 30          |
| 2.3 ELECTRONS IN CRYSTALLINE SOLIDS                                     | 33          |
| 2.3.1 Particle in a periodic potential: Bloch theorem                   | 36          |
| 2.4 OCCUPATION OF STATES:<br>DISTRIBUTION FUNCTION                      | 41          |

|          |                                                                                         |            |
|----------|-----------------------------------------------------------------------------------------|------------|
| 2.5      | METALS AND INSULATORS . . . . .                                                         | 43         |
| 2.5.1    | Electrons and Holes . . . . .                                                           | 45         |
| 2.6      | BANDSTRUCTURE OF SOME IMPORTANT SEMICONDUCTORS . . . . .                                | 46         |
| 2.6.1    | Direct and indirect semiconductors . . . . .                                            | 46         |
| 2.7      | MOBILE CARRIERS . . . . .                                                               | 50         |
| 2.7.1    | Mobile electrons in metals . . . . .                                                    | 50         |
| 2.7.2    | Electrons and holes in semiconductors . . . . .                                         | 55         |
| 2.8      | DOPING OF SEMICONDUCTORS . . . . .                                                      | 60         |
| 2.9      | DOPING IN POLAR MATERIALS . . . . .                                                     | 65         |
| 2.10     | TAILORING ELECTRONIC PROPERTIES . . . . .                                               | 73         |
| 2.10.1   | Electronic properties of alloys . . . . .                                               | 73         |
| 2.10.2   | Electronic properties of quantum wells . . . . .                                        | 73         |
| 2.11     | STRAINED HETEROSTRUCTURES . . . . .                                                     | 80         |
| 2.12     | DEFECT STATES IN SOLIDS . . . . .                                                       | 84         |
| 2.13     | TECHNOLOGY ISSUES . . . . .                                                             | 86         |
| 2.14     | PROBLEMS . . . . .                                                                      | 86         |
| 2.15     | FURTHER READING . . . . .                                                               | 90         |
| <b>3</b> | <b>Charge transport in materials</b>                                                    | <b>92</b>  |
| 3.1      | INTRODUCTION . . . . .                                                                  | 92         |
| 3.2      | CHARGE TRANSPORT: AN OVERVIEW . . . . .                                                 | 93         |
| 3.3      | TRANSPORT AND SCATTERING . . . . .                                                      | 94         |
| 3.3.1    | Quantum Mechanics and Scattering of electrons . . . . .                                 | 96         |
| 3.4      | TRANSPORT UNDER AN ELECTRIC FIELD . . . . .                                             | 104        |
| 3.4.1    | Velocity–electric field relations in semiconductors . . . . .                           | 104        |
| 3.5      | SOME IMPORTANT ISSUES IN TRANSPORT . . . . .                                            | 117        |
| 3.6      | CARRIER TRANSPORT BY DIFFUSION . . . . .                                                | 118        |
| 3.6.1    | Drift and diffusion transport: Einstein’s relation . . . . .                            | 121        |
| 3.7      | CHARGE INJECTION AND QUASI-FERMI LEVELS . . . . .                                       | 124        |
| 3.7.1    | Non-equilibrium Distributions . . . . .                                                 | 124        |
| 3.8      | CARRIER GENERATION AND RECOMBINATION . . . . .                                          | 125        |
| 3.8.1    | Optical Absorption and Emission in Semiconductors . . . . .                             | 128        |
| 3.8.2    | Schockley Read Hall Statistics . . . . .                                                | 133        |
| 3.9      | CURRENT CONTINUITY(The law of conservation of electrons and holes separately) . . . . . | 137        |
| 3.10     | PROBLEMS . . . . .                                                                      | 142        |
| 3.11     | FURTHER READING . . . . .                                                               | 145        |
| <b>4</b> | <b>Junctions in Semiconductors: <i>P-N</i> Diodes</b>                                   | <b>146</b> |
| 4.1      | Introduction . . . . .                                                                  | 146        |
| 4.2      | <i>P-N</i> JUNCTION IN EQUILIBRIUM . . . . .                                            | 146        |
| 4.3      | <i>P-N</i> DIODE UNDER BIAS . . . . .                                                   | 155        |
| 4.3.1    | Drift and Diffusion Currents in the Biased Diode . . . . .                              | 157        |

|          |                                                                                  |            |
|----------|----------------------------------------------------------------------------------|------------|
| 4.3.2    | Minority and Majority Currents in the $p$ - $n$ Diode . . . . .                  | 160        |
| 4.3.3    | Narrow Diode Current . . . . .                                                   | 160        |
| 4.4      | REAL DIODES: CONSEQUENCES OF DEFECTS AND CARRIER GENERATION . . . . .            | 163        |
| 4.4.1    | Generation-Recombination Currents . . . . .                                      | 164        |
| 4.5      | Reverse Bias Characteristics . . . . .                                           | 172        |
| 4.5.1    | First Observation . . . . .                                                      | 172        |
| 4.5.2    | Quasi Fermi Levels . . . . .                                                     | 175        |
| 4.6      | HIGH-VOLTAGE EFFECTS IN DIODES . . . . .                                         | 177        |
| 4.6.1    | Forward Bias: High Injection Region . . . . .                                    | 177        |
| 4.6.2    | Reverse Bias: Impact Ionization . . . . .                                        | 177        |
| 4.7      | Avalanche Breakdown in a $p$ - $n$ junction . . . . .                            | 178        |
| 4.7.1    | Reverse Bias: Zener Breakdown . . . . .                                          | 180        |
| 4.8      | DIODE APPLICATIONS: AN OVERVIEW . . . . .                                        | 183        |
| 4.8.1    | Applications of $p$ - $n$ diodes . . . . .                                       | 183        |
| 4.8.2    | The Solar Cell and Photodetector . . . . .                                       | 184        |
| 4.8.3    | The uses of diode non-linearity (Mixers, Multipliers, Power Detectors) . . . . . | 189        |
| 4.8.4    | Power Devices . . . . .                                                          | 191        |
| 4.9      | Light emitting diode (LED) . . . . .                                             | 193        |
| 4.9.1    | Emission Energy . . . . .                                                        | 193        |
| 4.9.2    | Carrier Injection and Spontaneous Emission . . . . .                             | 197        |
| 4.10     | PROBLEMS . . . . .                                                               | 204        |
| 4.11     | DESIGN PROBLEMS . . . . .                                                        | 211        |
| 4.12     | FURTHER READING . . . . .                                                        | 214        |
| <b>5</b> | <b>Semiconductor Junctions</b> . . . . .                                         | <b>216</b> |
| 5.1      | INTRODUCTION . . . . .                                                           | 216        |
| 5.2      | METAL INTERCONNECTS . . . . .                                                    | 216        |
| 5.3      | METAL SEMICONDUCTOR JUNCTION: SCHOTTKY BARRIER . . . . .                         | 219        |
| 5.3.1    | Schottky Barrier Height . . . . .                                                | 219        |
| 5.3.2    | Capacitance Voltage Characteristics . . . . .                                    | 223        |
| 5.3.3    | Current Flow across a Schottky Barrier: Thermionic Emission . . . . .            | 223        |
| 5.3.4    | Comparison of Schottky and $p$ - $n$ diodes . . . . .                            | 227        |
| 5.4      | METAL SEMICONDUCTOR JUNCTIONS FOR OHMIC CONTACTS . . . . .                       | 229        |
| 5.5      | INSULATOR-SEMICONDUCTOR JUNCTIONS . . . . .                                      | 230        |
| 5.5.1    | Insulator-Silicon . . . . .                                                      | 230        |
| 5.6      | SEMICONDUCTOR HETEROJUNCTIONS . . . . .                                          | 232        |
| 5.6.1    | Abrupt $p$ - $n$ heterojunction . . . . .                                        | 232        |
| 5.6.2    | Graded $p$ - $n$ heterojunction . . . . .                                        | 237        |
| 5.6.3    | Quasi-electric fields . . . . .                                                  | 238        |
| 5.7      | PROBLEMS . . . . .                                                               | 240        |

|          |                                                                       |            |
|----------|-----------------------------------------------------------------------|------------|
| 5.8      | FURTHER READING . . . . .                                             | 245        |
| <b>6</b> | <b>Bipolar Junction Transistors</b>                                   | <b>246</b> |
| 6.1      | INTRODUCTION . . . . .                                                | 246        |
| 6.2      | BIPOLAR TRANSISTOR: A CONCEPTUAL PICTURE . . . . .                    | 248        |
| 6.3      | STATIC CHARACTERISTICS: CURRENT-VOLTAGE<br>RELATION . . . . .         | 252        |
| 6.3.1    | Current Flow in a BJT . . . . .                                       | 253        |
| 6.3.2    | BJT Biasing in circuits . . . . .                                     | 259        |
| 6.3.3    | Current-Voltage: The Ebers-Moll Model . . . . .                       | 259        |
| 6.4      | DEVICE DESIGN AND DEVICE PERFORMANCE PARAMETERS . . . . .             | 262        |
| 6.5      | BJT DESIGN LIMITATIONS: NEED FOR BAND<br>TAILORING AND HBTs . . . . . | 265        |
| 6.5.1    | The Generalized Moll-Ross Relationship . . . . .                      | 269        |
| 6.5.2    | How much $\beta$ do we need? . . . . .                                | 270        |
| 6.6      | SECONDARY EFFECTS IN REAL DEVICES . . . . .                           | 274        |
| 6.6.1    | High Injection: The Kirk Effect . . . . .                             | 274        |
| 6.6.2    | High Injection: Thermal Effects . . . . .                             | 280        |
| 6.6.3    | Base Width Modulation: The Early Effect . . . . .                     | 280        |
| 6.6.4    | Drift Effects in the Base: Nonuniform Doping . . . . .                | 282        |
| 6.6.5    | Avalanche Breakdown . . . . .                                         | 283        |
| 6.6.6    | Low Injection Effects and Current Gain . . . . .                      | 284        |
| 6.6.7    | Current Crowding Effect . . . . .                                     | 284        |
| 6.7      | PROBLEMS . . . . .                                                    | 290        |
| 6.8      | DESIGN PROBLEMS . . . . .                                             | 301        |
| 6.9      | FURTHER READING . . . . .                                             | 303        |
| <b>7</b> | <b>Temporal Response Of Diodes and Bipolar Transistors</b>            | <b>304</b> |
| 7.1      | INTRODUCTION . . . . .                                                | 304        |
| 7.2      | MODULATION AND SWITCHING OF A <i>P-N</i> DIODE: AC RESPONSE . . . . . | 304        |
| 7.2.1    | Small-Signal Equivalent Circuit of a <i>p-n</i> Diode . . . . .       | 306        |
| 7.2.2    | Switching characteristics of diodes . . . . .                         | 312        |
| 7.3      | Temporal Response of a Schottky Diode . . . . .                       | 317        |
| 7.4      | BIPOLAR JUNCTION TRANSISTORS:<br>A CHARGE-CONTROL ANALYSIS . . . . .  | 318        |
| 7.4.1    | Junction Voltages at Saturation . . . . .                             | 324        |
| 7.5      | HIGH-FREQUENCY BEHAVIOR OF A BJT . . . . .                            | 325        |
| 7.5.1    | Bipolar Transistor Small-Signal Equivalent Circuit . . . . .          | 333        |
| 7.5.2    | Attenuation and Phase Shift of a Traveling Electron Wave . . . . .    | 336        |
| 7.5.3    | Small Signal Figures of Merit . . . . .                               | 340        |
| 7.6      | BIPOLAR TRANSISTORS: A TECHNOLOGY<br>ROADMAP . . . . .                | 345        |
| 7.6.1    | Si Bipolar Technology . . . . .                                       | 345        |
| 7.6.2    | Si-Based HBTs . . . . .                                               | 346        |

|          |                                                                                                       |            |
|----------|-------------------------------------------------------------------------------------------------------|------------|
| 7.6.3    | GaAs/AlGaAs HBTs . . . . .                                                                            | 347        |
| 7.6.4    | InGaAs/InAlAs and InGaAs/InP HBTs . . . . .                                                           | 347        |
| 7.7      | PROBLEMS . . . . .                                                                                    | 348        |
| 7.8      | DESIGN PROBLEMS . . . . .                                                                             | 355        |
| <b>8</b> | <b>Field Effect Transistors</b>                                                                       | <b>356</b> |
| 8.1      | INTRODUCTION . . . . .                                                                                | 356        |
| 8.2      | JFET AND MESFET: CHARGE CONTROL . . . . .                                                             | 356        |
| 8.3      | CURRENT-VOLTAGE CHARACTERISTICS . . . . .                                                             | 362        |
| 8.3.1    | The Ohmic Regime . . . . .                                                                            | 362        |
| 8.3.2    | A Nearly Universal Model for FET Behavior : The Saturation Regime . .                                 | 368        |
| 8.4      | HFETs: INTRODUCTION . . . . .                                                                         | 375        |
| 8.5      | CHARGE CONTROL MODEL FOR THE MODFET . . . . .                                                         | 378        |
| 8.5.1    | Modulation Efficiency . . . . .                                                                       | 385        |
| 8.6      | POLAR MATERIALS AND STRUCTURES . . . . .                                                              | 388        |
| 8.6.1    | Polar Materials . . . . .                                                                             | 388        |
| 8.6.2    | Polar HFET Structures . . . . .                                                                       | 394        |
| 8.7      | DESIGN ISSUES IN HFETs . . . . .                                                                      | 395        |
| 8.7.1    | $n^+$ Cap Layers . . . . .                                                                            | 395        |
| 8.7.2    | Maximizing 2DEG Conductivity . . . . .                                                                | 396        |
| 8.7.3    | Back-barriers to Substrate Injection . . . . .                                                        | 398        |
| 8.7.4    | Gate Recess Design . . . . .                                                                          | 400        |
| 8.7.5    | Field Plates . . . . .                                                                                | 401        |
| 8.7.6    | Comparison of two disparate material systems:<br>AlInAs/GaInAs and AlGaIn/GaN . . . . .               | 402        |
| 8.7.7    | Non-idealities in state-of-the-art transistors . . . . .                                              | 403        |
| 8.8      | SMALL AND LARGE SIGNAL ISSUES AND FIGURES OF MERIT . . . . .                                          | 411        |
| 8.8.1    | Small-Signal Characteristics . . . . .                                                                | 411        |
| 8.8.2    | Power-frequency limit . . . . .                                                                       | 415        |
| 8.8.3    | Classes of operation of transistor power amplifiers and necessary device<br>characteristics . . . . . | 416        |
| 8.9      | Implications on device technology and circuits . . . . .                                              | 421        |
| 8.10     | PROBLEMS . . . . .                                                                                    | 422        |
| 8.11     | DESIGN PROBLEMS . . . . .                                                                             | 431        |
| 8.12     | FURTHER READING . . . . .                                                                             | 432        |
| <b>9</b> | <b>Field Effect Transistors: MOSFET</b>                                                               | <b>433</b> |
| 9.1      | INTRODUCTION . . . . .                                                                                | 433        |
| 9.2      | MOSFET: DEVICES AND IMPACT . . . . .                                                                  | 434        |
| 9.3      | METAL-OXIDE-SEMICONDUCTOR CAPACITOR . . . . .                                                         | 437        |
| 9.4      | CAPACITANCE-VOLTAGE CHARACTERISTICS<br>OF THE MOS STRUCTURE . . . . .                                 | 448        |
| 9.5      | MOSFET OPERATION . . . . .                                                                            | 454        |
| 9.5.1    | Current-Voltage Characteristics . . . . .                                                             | 454        |

|           |                                                                            |            |
|-----------|----------------------------------------------------------------------------|------------|
| 9.5.2     | Substrate Bias Effects . . . . .                                           | 458        |
| 9.5.3     | Depletion and Enhancement MOSFETs . . . . .                                | 462        |
| 9.5.4     | Complementary MOSFETs . . . . .                                            | 464        |
| 9.6       | IMPORTANT ISSUES AND FUTURE CHALLENGES IN REAL MOSFETS . . . . .           | 466        |
| 9.6.1     | Subthreshold Conduction . . . . .                                          | 466        |
| 9.6.2     | Mobility Variation with Gate Bias . . . . .                                | 472        |
| 9.6.3     | Important Effects in Short-Channel MOSFETs . . . . .                       | 475        |
| 9.6.4     | Parasitic Bipolar Transistors and Latch-up in CMOS . . . . .               | 478        |
| 9.7       | SUMMARY . . . . .                                                          | 480        |
| 9.8       | PROBLEMS . . . . .                                                         | 481        |
| 9.9       | DESIGN PROBLEMS . . . . .                                                  | 485        |
| 9.10      | FURTHER READING . . . . .                                                  | 487        |
| <b>10</b> | <b>Coherent Transport and Mesoscopic Devices</b>                           | <b>489</b> |
| 10.1      | INTRODUCTION . . . . .                                                     | 489        |
| 10.2      | ZENER-BLOCH OSCILLATIONS . . . . .                                         | 491        |
| 10.3      | RESONANT TUNNELING . . . . .                                               | 492        |
| 10.4      | QUANTUM INTERFERENCE EFFECTS . . . . .                                     | 497        |
| 10.5      | MESOSCOPIC STRUCTURES . . . . .                                            | 500        |
| 10.5.1    | Conductance Fluctuations and Coherent Transport . . . . .                  | 500        |
| 10.5.2    | Coulomb Blockade Effects . . . . .                                         | 502        |
| 10.6      | MAGNETIC SEMICONDUCTORS<br>AND SPINTRONICS . . . . .                       | 505        |
| 10.7      | PROBLEMS . . . . .                                                         | 509        |
| 10.8      | Further Reading . . . . .                                                  | 510        |
| <b>A</b>  | <b>List of Symbols</b>                                                     | <b>511</b> |
| <b>B</b>  | <b>Boltzmann Transport Theory</b>                                          | <b>517</b> |
| B.1       | BOLTZMANN TRANSPORT EQUATION . . . . .                                     | 517        |
| B.1.1     | Diffusion-Induced Evolution of $f_{\mathbf{k}}(\mathbf{r})$ . . . . .      | 518        |
| B.1.2     | External Field-Induced Evolution of $f_{\mathbf{k}}(\mathbf{r})$ . . . . . | 519        |
| B.1.3     | Scattering-Induced Evolution of $f_{\mathbf{k}}(\mathbf{r})$ . . . . .     | 519        |
| B.2       | AVERAGING PROCEDURES . . . . .                                             | 526        |
| <b>C</b>  | <b>Density of States</b>                                                   | <b>529</b> |
| <b>D</b>  | <b>Important Properties of Semiconductors</b>                              | <b>535</b> |
| <b>E</b>  | <b>Beyond the Depletion Approximation</b>                                  | <b>544</b> |
| <b>F</b>  | <b>Design of Graded Heterojunctions for Bipolar Transistors</b>            | <b>548</b> |
|           | <b>Index</b>                                                               | <b>552</b> |

# INTRODUCTION

The pace of semiconductor materials and device development has been staggering, and the impact on human society monumental. Leading this advance has been the development of the silicon-based MOSFET device and its continuous high level of integration. Moore's Law (shown in figure .1), which predicts the doubling of device density every 18 months, has been the governing maxim of the industry. Sustaining Moore's Law has required:

- The development of lithography tools to achieve the 45 nm gate length MOSFETs released into production in 2006
- The continuous scaling of silicon wafers to 12 inch diameters (2005) and 15 inch in the future to enable large chip yields per wafer
- Tremendous advances in interconnect technology
- Device innovations to continuously maintain charge control and low gate leakage as the oxide thickness is scaled down along with the gate length

Though most of the chip and dollar volume of the industry has been driven by Si-based CMOS architecture, there have been critical advances made in other semiconductor technologies. The ability to grow epitaxial layers in a controlled fashion, initially by Liquid Phase Epitaxy (LPE) and Vapor Phase Epitaxy (VPE) and currently by Metalorganic Vapor Phase Epitaxy (MOVPE) and Molecular Beam Epitaxy (MBE), has enabled the compound semiconductor industry to mature into a small but critical component of the total space. The impact has been felt in both the electronics and photonics arenas. In the former, development of the Heterojunction Field Effect Transistor (HFET) and the Heterojunction Bipolar Transistor (HBT) has had a large impact on analog and mixed signal applications. In the low noise receiver area, GaAs and InP based HFETs are the preferred technology. The GaAs-based HBT is preferred for power amplifiers in cellular phones. The Si/SiGe HBT is being actively used in mixed signal applications such as A/D converters and in BiCMOS implementations.

In the optical arena, the development of Light Emitting Diodes (LEDs), lasers, and detectors has been profound. LEDs are used in prolific applications such as signage displays and remote controls as well as in communication devices. The advent of GaN-based LEDs has raised the possibility of a revolutionary advance in lighting with the emergence of solid-state lighting.

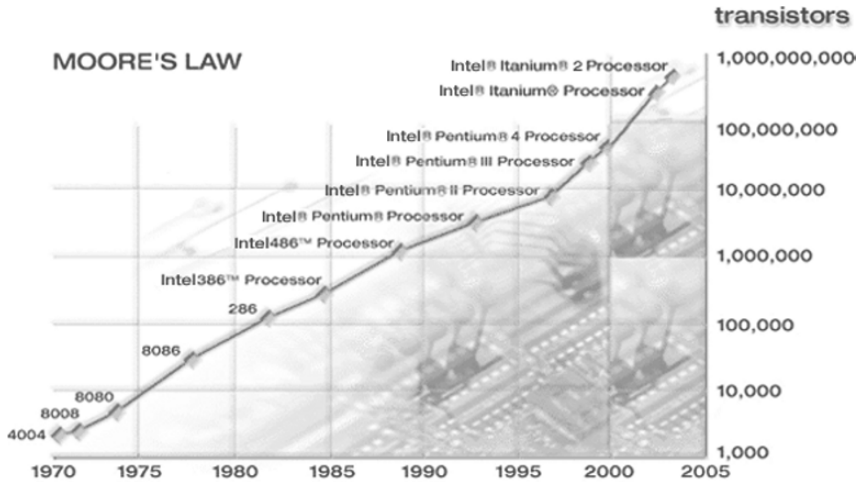


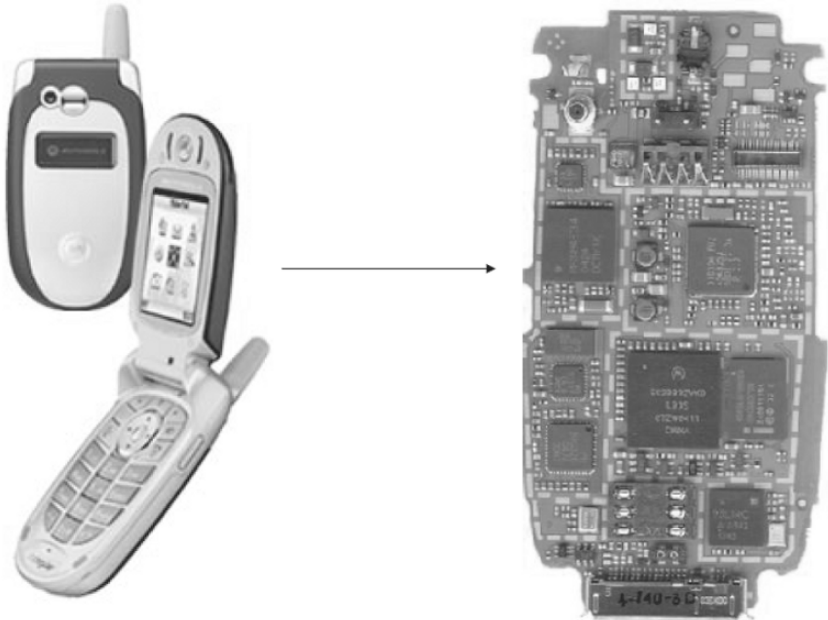
Figure .1: Illustration of Moore's Law.

Lasers and detectors have been the enabling elements in optical communications. Lasers have also enabled entertainment devices such as the DVD.

The continuous expansion of the material and device tool set has enabled system designers to choose the correct technology for the application, resulting in phenomenal advances at the system level. This is best understood by studying a commercial widespread system - the cellular phone. Consider the Motorola V551 phone, illustrated in figure .2. The key components of the transmit/receive chain in any radio architecture are the switch, filter, modulator/demodulator, LNA, mixer, gain blocks, and power amplifier. Integrating the different chips into a total radio solution places varied specifications on the different chips used to achieve the radio solution. In turn, these specifications drive the selection of the active device and process technology that is used to implement the functionality of the particular chip.

As an example of this technology selection process, consider the POLARIS total radio solution from RFMD, which is a highly integrated transceiver that performs all functions of the handset radio section, operating under GSM/GPRS/EDGE standards. The POLARIS chipset consists of the following functional blocks, shown in figure .3:

1. The RF 2722 quad-band RF receiver.
2. The RF 3146 POWER STAR PA module with integrated power control.
3. The RF 6001 digital filter, fractional-N PLL, modulator, and power amplifier ramp control.



Motorola V551

Figure .2: The Motoral V551 cellular phone. Picture courtesy of A. Upton, R. Vetury, and J. Shealy, RFMD.

The RF 2722 fulfills the functional requirement of a quad band LNA and mixer. It includes a VCO and supports very low IF (VLIF) and direct conversion receive (DCR) architectures, thus eliminating the need for IF SAW and RF interstage filters. The complexity of the circuit architecture needed and the noise and linearity requirements placed by the LNA and mixer make the technology of choice SiGe-BiCMOS.

The RF 3146 fulfills the functional requirement of a power amplifier. It includes considerable power control circuitry and can be driven from the DAC output, thus eliminating the need for directional couplers, detector diodes, and other power control circuitry. GaAs HBT technology is chosen for this component in order to achieve the optimum combination of high power, high PAE, and excellent linearity requirements at the frequency of operation.

The RF 6001 fulfills the functional requirements of a synthesizer and signal processor. To achieve the optimum combination of low cost, high levels of integration, and low power consumption, Si CMOS technology is chosen for this component.

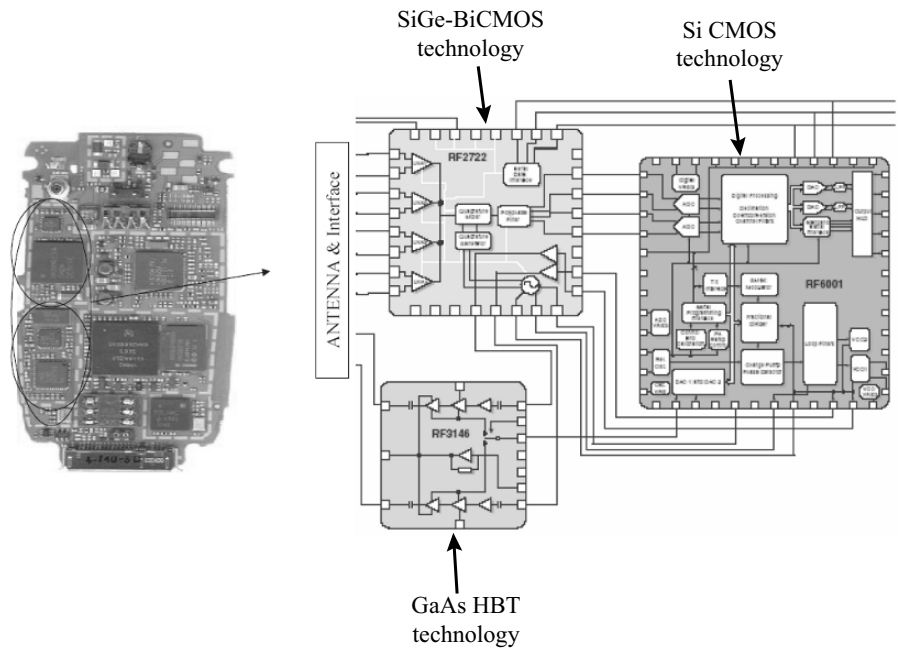


Figure .3: The POLARIS total radio solution from RFMD. Picture courtesy of A. Upton, R. Vetry, and J. Shealy, RFMD.

So what does the future hold for semiconductor based device development? There are brick walls facing the conventional scaling of CMOS circuits. Beyond the year 2012 and the 18 nm node, several of the pathways to continued scaling are not obvious. Also, the power dissipation in the chip threatens to set a thermal limit to the size and the speed of processors in the future. This is best illustrated in figure .4, where it is clear that today's chips seem hotter than a hot plate, and chips of the future in a tongue-in-cheek prediction may rival the sun's surface (obviously impossible). Hence now is the time for all of us to rethink the conventional CMOS scaling paradigm and consider what new pathways may open up. Could compound semiconductors, with their high electron mobilities and velocities, play a role in achieving high clock speeds and reduced power levels? Could large bandgap materials such as Gallium Nitride play a role in applications where the operating temperature is continuously rising? Are there completely new devices such as Carbon Nanotubes (CNTs) which operate in the ballistic regime of electron and hole transport that can emerge as the dark horse in future complementary circuits? Or is molecular electronics, the use of the electronic states of the molecule to achieve computation, the

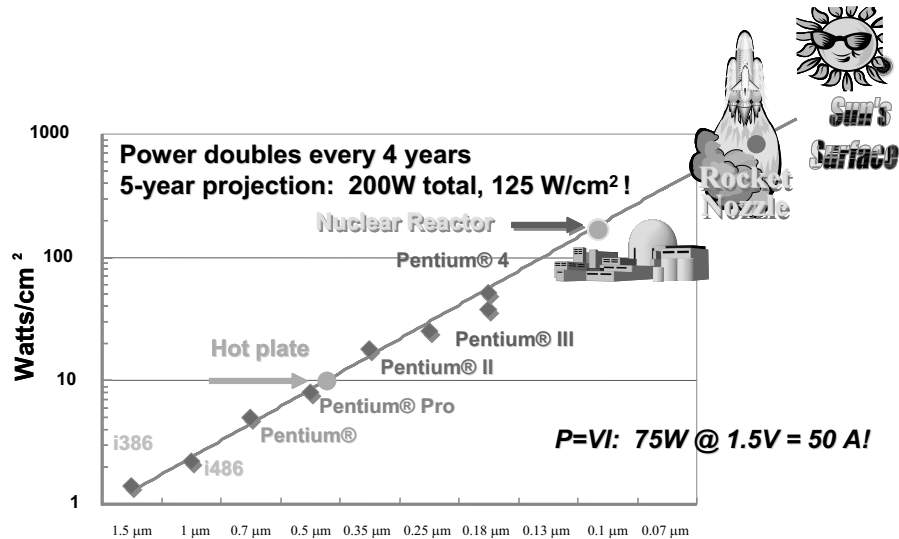


Figure .4: Chip power density is increasing exponentially with time.

answer? Is the control of electron spin rather than the total charge in the channel of the device (the emerging field of spintronics) the holy grail? Are architectures based on single electron transistors a high density, low power alternative?

The future is murky, and we as scientists and engineers have to help clarify it. This book seeks to provide an understanding of the materials, devices, and technology of the various alternatives being considered, with detail appropriate to the maturity of the technology. A bias towards compound semiconductors is obvious, as Si-based devices have been exclusively addressed over the years in various forms. We hope that this book serves a function to academics teaching course materials, engineers and researchers in the field tackling the murky future, and today's graduate students who will be the great engineers of tomorrow.

# Chapter 1

# STRUCTURAL PROPERTIES OF SEMICONDUCTORS

## 1.1 INTRODUCTION

In this text we will be exploring state of the art electronic devices that drive modern information technology. Essentially all of these devices are based on semiconductors. Semiconductor structures have also provided the stages for exploring questions of fundamental physics. As technology advances the number of semiconductors that are used in technology steadily increases. Indeed many innovations have arisen as a result of using new materials and their heterostructures. Thus while Si, GaAs and InP have been most widely used, other materials like InAs, GaN, InN etc. are finding important uses as well. It is important to recognize that the ability to examine fundamental physics issues and to use semiconductors in state of the art device technologies depends critically on the purity and perfection of the semiconductor crystal.

Semiconductor structures can operate at their potential only if they can be grown with a high degree of crystallinity and if impurities and defects can be controlled. For high structural quality it is essential that a high quality substrate be available. This requires growth of bulk crystals which are then sliced and polished to allow epitaxial growth of thin semiconductor regions including heterostructures.

In this chapter we will discuss important semiconductor crystal structures. We also discuss strained lattice structures and the strain tensor for such crystals. Strained epitaxy and its resultant consequences are now widely exploited in semiconductor physics. High speed SiGe devices are based on strained systems as are InGaAs and AlGaIn/GaN microwave devices.

We will start with some general properties of crystalline materials and then discuss some specific crystal structures important for semiconductors.

## 1.2 CRYSTAL STRUCTURE

As noted above high performance semiconductor devices are based on crystalline materials. Crystals are periodic structures made up of identical building blocks. While in “natural” crystals the crystalline symmetry is fixed by nature, new advances in crystal growth techniques are allowing scientists to produce artificial crystals with modified crystalline structure. These advances depend upon being able to place atomic layers with exact precision and control during growth, leading to “low dimensional systems”. To define the crystal structure, two important concepts are introduced. The lattice represents a set of points in space forming a periodic structure. The lattice is by itself a mathematical abstraction. A building block of atoms called the basis is then attached to each lattice point yielding the physical crystal structure.

To define a lattice one defines three vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ , such that any lattice point  $\mathbf{R}'$  can be obtained from any other lattice point  $\mathbf{R}$  by a translation

$$\mathbf{R}' = \mathbf{R} + m_1\mathbf{a}_1 + m_2\mathbf{a}_2 + m_3\mathbf{a}_3 \quad (1.2.1)$$

where  $m_1$ ,  $m_2$ ,  $m_3$  are integers. Such a lattice is called a Bravais lattice. The crystalline structure is now produced by attaching the basis to each of these lattice points.

$$\text{lattice} + \text{basis} = \text{crystal structure} \quad (1.2.2)$$

The translation vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ , and  $\mathbf{a}_3$  are called primitive if the volume of the cell formed by them is the smallest possible. There is no unique way to choose the primitive vectors. It is possible to define more than one set of primitive vectors for a given lattice, and often the choice depends upon convenience. The volume cell enclosed by the primitive vectors is called the primitive unit cell.

Because of the periodicity of a lattice, it is useful to define the symmetry of the structure. The symmetry is defined via a set of point group operations which involve a set of operations applied around a point. The operations involve rotation, reflection and inversion. The symmetry plays a very important role in the electronic properties of the crystals. For example, the inversion symmetry is extremely important and many physical properties of semiconductors are tied to the absence of this symmetry. As will be clear later, in the diamond structure (Si, Ge, C, etc.), inversion symmetry is present, while in the Zinc Blende structure (GaAs, AlAs, InAs, etc.), it is absent. Because of this lack of inversion symmetry, these semiconductors are piezoelectric, i.e., when they are strained an electric potential is developed across the opposite faces of the crystal. In crystals with inversion symmetry, where the two faces are identical, this is not possible.

### 1.2.1 Basic Lattice Types

The various kinds of lattice structures possible in nature are described by the symmetry group that describes their properties. Rotation is one of the important symmetry groups. Lattices can be found which have a rotation symmetry of  $2\pi$ ,  $\frac{2\pi}{2}$ ,  $\frac{2\pi}{3}$ ,  $\frac{2\pi}{4}$ ,  $\frac{2\pi}{6}$ . The rotation symmetries are

denoted by 1, 2, 3, 4, and 6. No other rotation axes exist; e.g.,  $\frac{2\pi}{5}$  or  $\frac{2\pi}{7}$  are not allowed because such a structure could not fill up an infinite space.

There are 14 types of lattices in 3D. These lattice classes are defined by the relationships between the primitive vectors  $a_1$ ,  $a_2$ , and  $a_3$ , and the angles  $\alpha$ ,  $\beta$ , and  $\gamma$  between them. We will focus on the cubic and hexagonal lattices which underly the structure taken by all semiconductors.

There are 3 kinds of cubic lattices: simple cubic, body centered cubic, and face centered cubic.

**Simple cubic:** The simple cubic lattice shown in figure 1.1 is generated by the primitive vectors

$$a\mathbf{x}, a\mathbf{y}, a\mathbf{z} \quad (1.2.3)$$

where the  $\mathbf{x}$ ,  $\mathbf{y}$ ,  $\mathbf{z}$  are unit vectors.

**Body-centered cubic :** The bcc lattice shown in figure 1.2 can be generated from the simple cubic structure by placing a lattice point at the center of the cube. If  $\hat{\mathbf{x}}$ ,  $\hat{\mathbf{y}}$ , and  $\hat{\mathbf{z}}$  are three orthogonal unit vectors, then a set of primitive vectors for the body-centered cubic lattice could be

$$a_1 = a\hat{\mathbf{x}}, a_2 = a\hat{\mathbf{y}}, a_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \quad (1.2.4)$$

A more symmetric set for the bcc lattice is

$$a_1 = \frac{a}{2}(\hat{\mathbf{y}} + \hat{\mathbf{z}} - \hat{\mathbf{x}}), a_2 = \frac{a}{2}(\hat{\mathbf{z}} + \hat{\mathbf{x}} - \hat{\mathbf{y}}), a_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}) \quad (1.2.5)$$

**Face Centered Cubic:** Another equally important lattice for semiconductors is the face-centered cubic (fcc) Bravais lattice shown in figure 1.3. To construct the face-centered cubic Bravais lattice add to the simple cubic lattice an additional point in the center of each square face. This form of packing is called close-packed.

A symmetric set of primitive vectors for the face-centered cubic lattice (see figure 1.3) is

$$a_1 = \frac{a}{2}(\hat{\mathbf{y}} + \hat{\mathbf{z}}), a_2 = \frac{a}{2}(\hat{\mathbf{z}} + \hat{\mathbf{x}}), a_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}}) \quad (1.2.6)$$

The face-centered cubic and body-centered cubic Bravais lattices are of great importance, since an enormous variety of solids crystallize in these forms with an atom (or ion) at each lattice site. Essentially all semiconductors of interest for electronics and optoelectronics have a close-packed structure, either fcc or Hexagonal Close Pack(HCP) as discussed below.

## 1.2.2 Basic Crystal Structures

### Diamond and Zinc Blende Structures

Most semiconductors of interest for electronics and optoelectronics have an underlying fcc lattice, with two atoms per basis. The coordinates of the two basis atoms are

$$(000) \text{ and } \left(\frac{a}{4}, \frac{a}{4}, \frac{a}{4}\right) \quad (1.2.7)$$

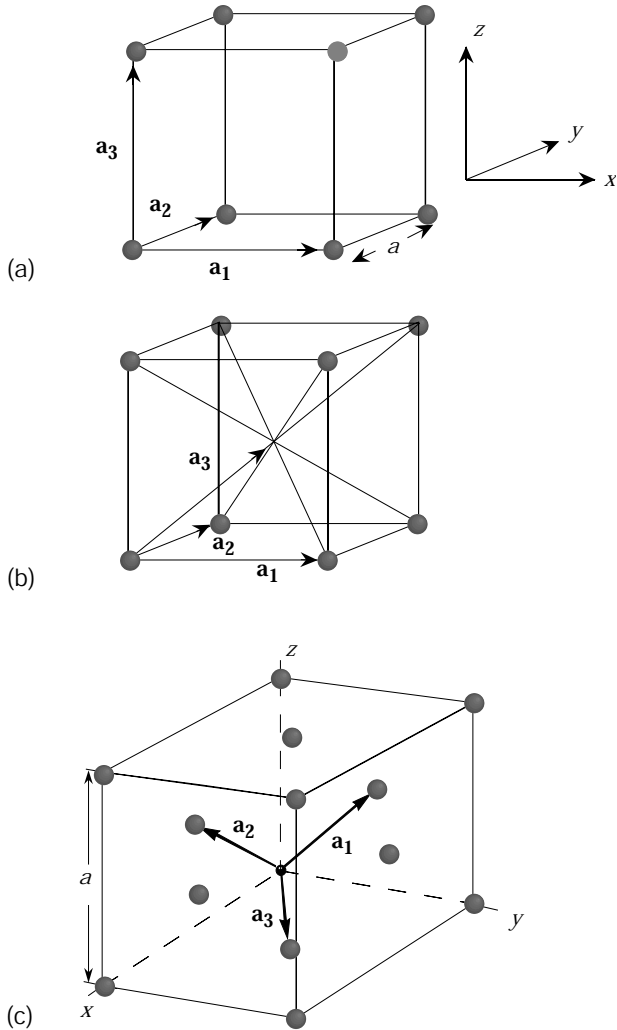


Figure 1.1: A simple cubic lattice showing the primitive vectors. The crystal is produced by repeating the cubic cell through space.

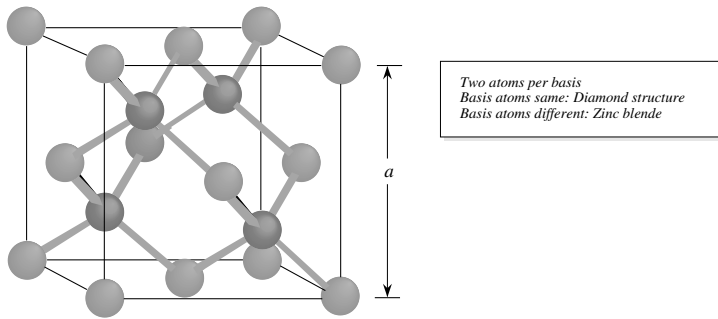


Figure 1.2: The body centered cubic lattice along with a choice of primitive vectors.

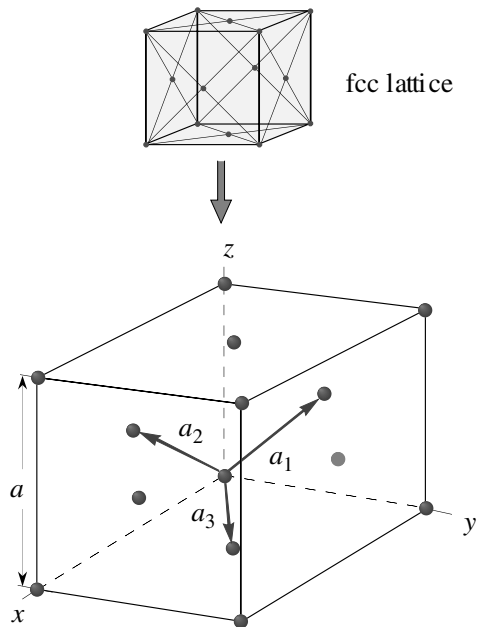


Figure 1.3: Primitive basis vectors for the face centered cubic lattice.

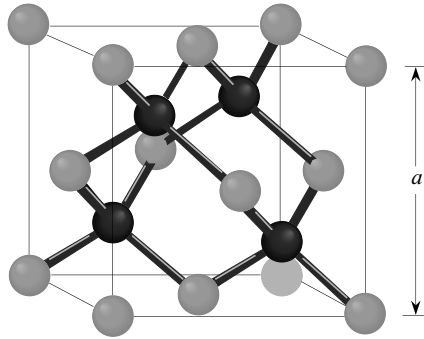


Figure 1.4: The zinc blende crystal structure. The structure consists of the interpenetrating fcc lattices, one displaced from the other by a distance  $(\frac{a}{4}, \frac{a}{4}, \frac{a}{4})$  along the body diagonal. The underlying Bravais lattice is fcc with a two atom basis. The positions of the two atoms is (000) and  $(\frac{a}{4}, \frac{a}{4}, \frac{a}{4})$ .

Since each atom lies on its own fcc lattice, such a two atom basis structure may be thought of as two inter-penetrating fcc lattices, one displaced from the other by a translation along the body diagonal direction  $(\frac{a}{4}, \frac{a}{4}, \frac{a}{4})$ .

Figure 1.4 shows this important structure. If the two atoms of the basis are identical, the structure is called diamond. Semiconductors such as Si, Ge, C, etc., fall in this category. If the two atoms are different, the structure is called the Zinc Blende structure. Semiconductors such as GaAs, AlAs, CdS, etc., fall in this category. Semiconductors with diamond structure are often called elemental semiconductors, while the Zinc Blende semiconductors are called compound semiconductors. The compound semiconductors are also denoted by the position of the atoms in the periodic chart, e.g., GaAs, AlAs, InP are called III-V (three-five) semiconductors while CdS, HgTe, CdTe, etc., are called II-VI (two-six) semiconductors.

**Hexagonal Close Pack Structure** The hexagonal close pack (hcp) structure is an important lattice structure and many semiconductors such as BN, AlN, GaN, SiC, etc., also have this underlying lattice (with a two-atom basis). The hcp structure is formed as shown in figure 1.5a. Imagine that a close-packed layer of spheres is formed. Each sphere touches six other spheres, leaving cavities, as shown in figure 1.5. A second close-packed layer of spheres is placed on top of the first one so that the second layer sphere centers are in the cavities formed by the first layer. The third layer of close-packed spheres can now be placed so that center of the spheres do not fall on the center of the starting spheres (left side of figure 1.5a) or coincide with the centers of the starting spheres (right side of figure 1.5). These two sequences, when repeated, produce the fcc and hcp lattices.

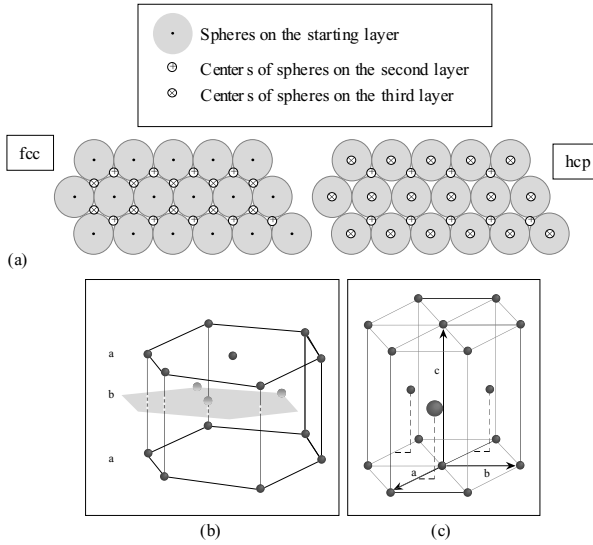


Figure 1.5: (a) A schematic of how the fcc and hcp lattices are formed by close packing of spheres. (b) The hcp structure is produced by two interpenetrating hexagonal lattices with a displacement discussed in the text. (c) Arrangement of lattice points on an hcp lattice.

In figure 1.5b and figure 1.5c we show the detailed positions of the lattice points in the hcp lattice. The three lattice vectors  $a_1$ ,  $a_2$ ,  $a_3$  are shown as  $a$ ,  $b$ ,  $c$ . The vector  $a_3$  is denoted by  $c$  and the term  $c$ -axis refers to the orientation of  $a_3$ . The hexagonal planes are displaced from each other by  $a_1/3 + a_2/3 + a_3/2$ . In an ideal structure, if  $|a| = |a_1| = |a_2|$ ,

$$\frac{c}{a} = \sqrt{\frac{8}{3}} \quad (1.2.8)$$

In table 1.1 we show the structural and some important electronic properties of some important semiconductors. Note that two or more semiconductors are randomly mixed to produce an alloy,  $A_xB_{1-x}$ , the lattice constant of the alloy is given by Vegard's law according to which the alloy lattice constant is the weighted mean of the lattice constants of the individual components

$$a_{\text{alloy}} = xa_A + (1-x)a_B \quad (1.2.9)$$

### 1.2.3 Notation to Denote Planes and Points in a Lattice: Miller Indices

To represent the directions and planes in a crystalline structure an agreed upon scheme is used. The planes or directions are denoted by a series of integers called the Miller indices.

it has a positive charge. It therefore responds to external electric and magnetic fields  $\mathcal{E}$  and  $\mathbf{B}$ , respectively, according to the equation of motion

$$\hbar \frac{d\mathbf{k}_h}{dt} = e [\mathcal{E} + \mathbf{v}_h \times \mathbf{B}] \quad (2.5.4)$$

where  $\hbar\mathbf{k}_h$  and  $\mathbf{v}_h$  are the momentum and velocity of the hole.

Thus the equation of motion of holes is that of particles with a positive charge  $e$ . The mass of the hole has a positive value, although the electron mass in its valence band is negative. When we discuss the valence band properties, we refer to holes. This is because in the valence band only the missing electrons or holes lead to charge transport and current flow.

## 2.6 BANDSTRUCTURE OF SOME IMPORTANT SEMICONDUCTORS

In this section we will examine the band structure near the band edges for several important materials. To represent the bandstructure on a figure that is two-dimensional, we draw the  $E$ - $k$  diagram in several panels where  $k$  goes from zero to its maximum value along the (100) direction or the (111) direction, etc within the Brillouin zone. As shown in figure 2.6 for the fcc lattice, the maximum  $k$ -value along the (100) direction is  $2\pi/a(1, 0, 0)$ . This point is called the  $X$ -point and there are five other equivalent points, due to the cubic symmetry of the lattice. Similarly, along the (111) direction, the maximum  $k$ -point is  $\pi/a(1, 1, 1)$  and seven other similar points. This point is called the  $L$ -point. Thus we commonly display the  $E$ - $k$  diagram with  $k$  going from the origin (called the  $\Gamma$ -point) to the  $X$ -point and from the origin to the  $L$ -point.

### 2.6.1 Direct and indirect semiconductors

Two types of band structures arise in semiconductors- direct and indirect. The top of the valence band of most semiconductors occurs at effective momentum equal to zero. A typical bandstructure of a semiconductor near the top of the valence band is shown in figure 2.11. We notice the presence of three bands near the valence band edge. These curves or bands are labeled I, II, and III in the figure and are called the heavy hole (HH), light hole (LH), and the split off hole bands.

The bottom of the conduction band in some semiconductors occurs at  $k = 0$ . Such semiconductors are called direct bandgap materials. Semiconductors, such as GaAs, InP, GaN, InN, etc., are direct bandgap semiconductors. In other semiconductors, the bottom of the conduction band does not occur at the  $k = 0$  point, but at certain other points. Such semiconductors are called indirect semiconductors. Examples are Si, Ge, AlAs, etc.

Due to the law of momentum conservation, direct gap materials have a strong interaction with light. Indirect gap materials have a relatively weak interaction with electrons.

When the band edges are at  $k = 0$  it is possible to represent the bandstructure by a simple relation of the form

$$E(\mathbf{k}) = E_c + \frac{\hbar^2 k^2}{2m^*} \quad (2.6.1)$$

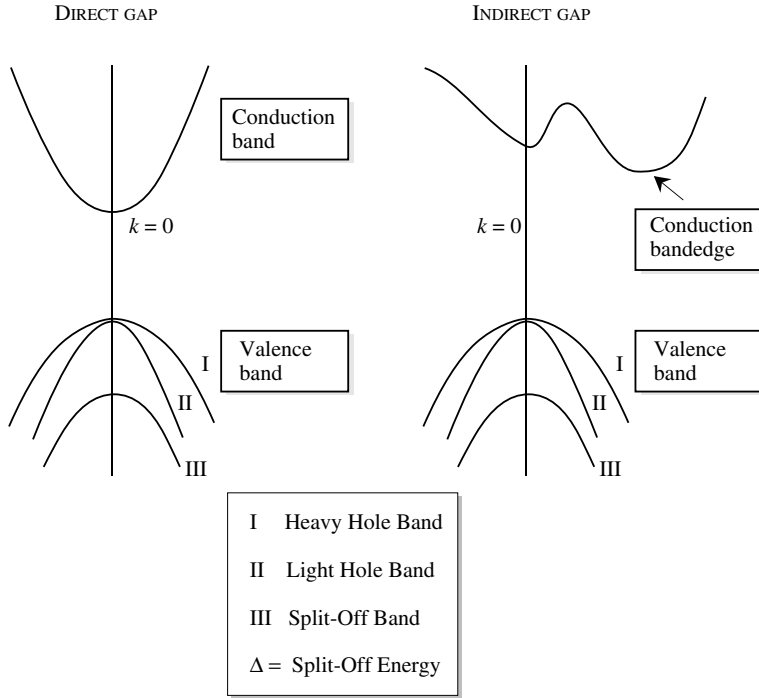


Figure 2.11: Schematic of the valence band, direct bandgap, and indirect bandgap conduction bands. The curves I, II, III in the valence band are called heavy hole, light hole, and split-off hole states, respectively.

where  $E_c$  is the conduction bandedge, and the bandstructure is a simple parabola. The equation for the  $E$ - $k$  relation looks very much like that of an electron in free space as noted in the previous section.

### Silicon

The most important semiconductor is silicon. Silicon has an indirect bandgap as shown in figure 2.12. The bottom of the conduction band in Si is at point  $(\sim (2\pi/a)(0.85, 0.0))$ ; i.e., close to the  $X$ -point. There are six degenerate  $X$ -points and, consequently, six conduction bandedge valleys. The near bandedge bandstructure can be represented by ellipsoids of energy with simple  $E$  vs.  $k$  relations of the form (for examples for the  $[100]$  valley)

$$E(\mathbf{k}) = \frac{\hbar^2 k_x^2}{2m_t^*} + \frac{\hbar^2 (k_y^2 + k_z^2)}{2m_l^*} \quad (2.6.2)$$

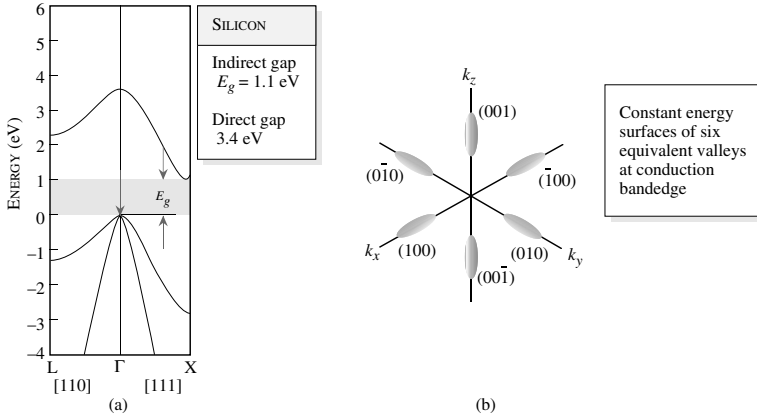


Figure 2.12: (a) Bandstructure of Si. (b) Constant energy ellipsoids for the Si conduction band. There are six equivalent valley in Si at the bandedge.

where we have two masses, the longitudinal and transverse. The constant energy surfaces of Si are ellipsoids according to Eq. 2.6.2. The six surfaces are shown in figure 2.12

The longitudinal electron mass  $m_l^*$  is approximately  $0.98 m_0$ , while the transverse mass is approximately  $0.19 m_0$ .

The next valley in the conduction band is the  $L$ -point valley, which is about 1.1 eV above the bandedge. Above this is the  $\Gamma$ -point edge. Due to the six-fold degeneracy of the conduction bandedge, the electron transport in Si is quite poor because of the very large density of states near the bandedge, leading to a high scattering rate in transport.

### GaAs

GaAs is a direct gap material with small electron effective mass. The near bandedge bandstructure of GaAs is shown in figure 2.13. The bandstructure can be represented by the relation (referenced to  $E_c$ )

$$E = \frac{\hbar^2 k^2}{2m^*} \quad (2.6.3)$$

with  $m^* = 0.067m_0$ . A better relationship is the non-parabolic approximation

$$E(1 + \alpha E) = \frac{\hbar^2 k^2}{2m^*} \quad (2.6.4)$$

with  $\alpha = 0.67 \text{ eV}^{-1}$ .

For high electric field transport, it is important to note that the valleys above  $\Gamma$ -point are the  $L$ -valleys. There are eight  $L$ -points, but, since half of them are connected by a reciprocal lattice vector, there are four valleys. The separation  $\Delta E_{\Gamma L}$  between the  $\Gamma$ - and  $L$ - minima is 0.29 eV.

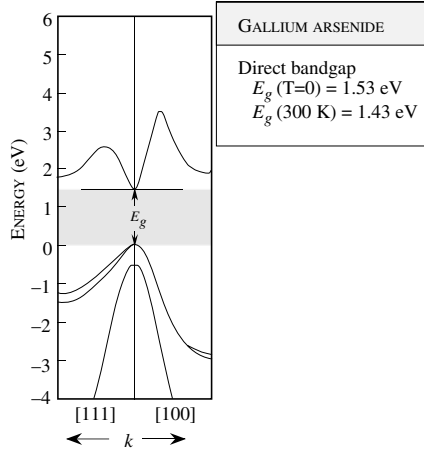


Figure 2.13: Bandstructure of GaAs. The bandgap at 0 K is 1.51 eV and at 300 K it is 1.43 eV. The bottom of the conduction band is at  $k = (0, 0, 0)$ , i.e., the  $\Gamma$ -point. The upper conduction band valleys are at the  $L$ -point.

The  $L$ -valley has a much larger effective mass than the  $\Gamma$ -valley. For GaAs,  $m_L^* \sim 0.25m_0$ . This difference in masses is extremely important for high electric field transport as will be discussed in the next chapter.

The valence band of GaAs has the standard HH, LH, and SO bands. Due to the large spin–orbit splitting, for most purposes the SO band does not play any role in electronic properties.

The bandstructures of Ge and AlAs, two other important semiconductors, are shown in figure 2.14, along with brief comments about their important properties.

### InN, GaN, and AlN

The III–V nitride family of GaN, InN, and AlN have become quite important due to progress in the ability to grow the semiconductor. These materials are typically grown with a wurtzite structure, and have bandgaps ranging from  $\sim 1.0 \text{ eV}$  to over  $6.0 \text{ eV}$ . This large bangap is very useful for short wavelength light emitters and high power electronics. In figure 2.15 we show the bandstructure for nitrides.

It is important to note is that the bandgap of semiconductors generally decreases as temperature increases. The bandgap of GaAs, for example, is  $1.51 \text{ eV}$  at  $T = 0\text{K}$  and  $1.43 \text{ eV}$  at room temperature. In table 2.1 we show the temperature dependence of bandgaps of several semiconductors.

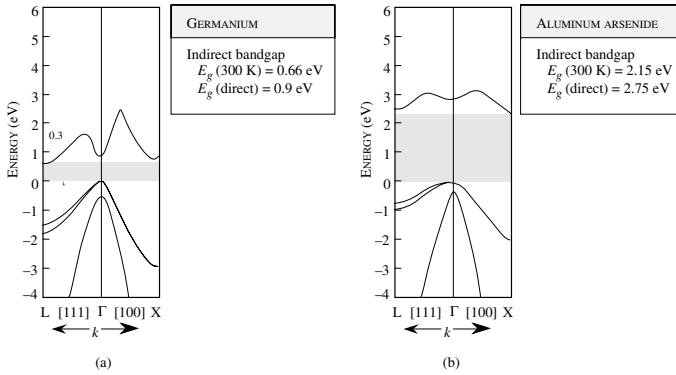


Figure 2.14: (a) Bandstructure of Ge and AlAs.

## 2.7 MOBILE CARRIERS

From our brief discussion of metals and semiconductors in Section 2.5, we see that in a metal current flows because of the electrons present in the highest (partially) filled band. As shown schematically in figure 2.16a. The density of such electrons is very high ( $\sim 10^{23} \text{ cm}^{-3}$ ). In a semiconductor, in contrast, no current flows if the valence band is filled with electrons and the conduction band is empty of electrons. However, if somehow empty states or holes are created in the valence band by removing electrons, current can flow through the holes. Similarly, if electrons are placed in the conduction band, these electrons can carry current. This is shown schematically in figure 2.16b. If the density of electrons in the conduction band is  $n$  and that of holes in the valence band is  $p$ , the total mobile carrier density is  $n + p$ .

### 2.7.1 Mobile electrons in metals

In a metal, we have a series of filled bands and a partially filled band called the conduction band. The filled bands are inert as far as electrical and optical properties of metals are concerned. The conduction band of metals can be assumed to be described by the parabolic energy-momentum relation

$$E(k) = E_c + \frac{\hbar^2 k^2}{2m_0} \quad (2.7.1)$$

Note that we have used an effective mass equal to the free electrons mass. This is a reasonable approximation for metals. The large electron density in the band “screens” out the background potential and the electron effective mass is quite close to the free space value.

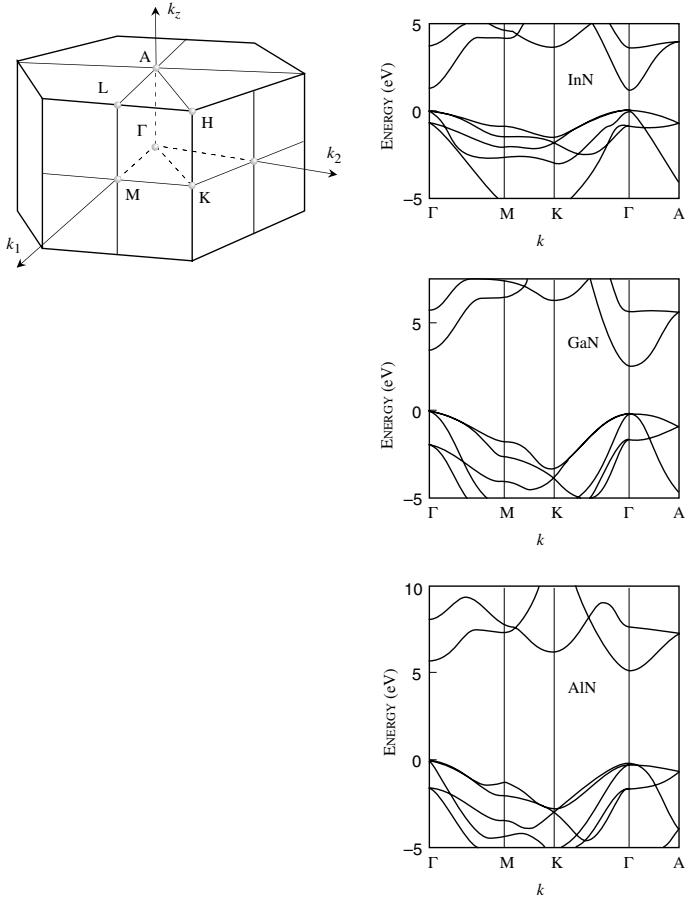


Figure 2.15: Bandstructure of InN, GaN and AlN.

The electron density in the conduction band of a metal is related to the Fermi level by the relation

$$n = \int_{E_c}^{\infty} \frac{\sqrt{2}m_0^{3/2}}{\pi^2\hbar^3} \frac{E^{1/2}dE}{\exp\left(\frac{E-E_F}{k_B T}\right) + 1} \quad (2.7.2)$$

| Compound | Type of bandgap | Experimental bandgap $E_G$ (eV) |       | Temperature dependence of bandgap $E_G(T)$ (eV) |
|----------|-----------------|---------------------------------|-------|-------------------------------------------------|
|          |                 | 0 K                             | 300 K |                                                 |
| AlP      | Indirect        | 2.52                            | 2.45  | $2.52 - 3.18 \times 10^{-4} T^2 / (T + 588)$    |
| AlAs     | Indirect        | 2.239                           | 2.163 | $2.239 - 6.0 \times 10^{-4} T^2 / (T + 408)$    |
| AlSb     | Indirect        | 1.687                           | 1.58  | $1.687 - 4.97 \times 10^{-4} T^2 / (T + 213)$   |
| GaP      | Indirect        | 2.338                           | 2.261 | $2.338 - 5.771 \times 10^{-4} T^2 / (T + 372)$  |
| GaAs     | Direct          | 1.519                           | 1.424 | $1.519 - 5.405 \times 10^{-4} T^2 / (T + 204)$  |
| GaSb     | Direct          | 0.810                           | 0.726 | $0.810 - 3.78 \times 10^{-4} T^2 / (T + 94)$    |
| InP      | Direct          | 1.421                           | 1.351 | $1.421 - 3.63 \times 10^{-4} T^2 / (T + 162)$   |
| InAs     | Direct          | 0.420                           | 0.360 | $0.420 - 2.50 \times 10^{-4} T^2 / (T + 75)$    |
| InSb     | Direct          | 0.236                           | 0.172 | $0.236 - 2.99 \times 10^{-4} T^2 / (T + 140)$   |

Table 2.1: Bandgaps of binary III–V compounds (From Casey and Panish, 1978).

This integral is particularly simple to evaluate as 0 K, since, at this temperature

$$\frac{1}{\exp\left(\frac{E-E_F}{k_B T}\right) + 1} = \begin{cases} 1 & \text{if } E \leq E_F \\ 0 & \text{otherwise} \end{cases}$$

this gives

$$n = \int_{E_C}^{E_F} N(E) dE$$

We then have

$$\begin{aligned} n &= \frac{\sqrt{2}m_0^{3/2}}{\pi^2 \hbar^3} \int_{E_C}^{E_F} (E - E_C)^{1/2} dE \\ &= \frac{2\sqrt{2}m_0^{3/2}}{3\pi^2 \hbar^3} (E_F - E_C)^{3/2} \end{aligned}$$

or

$$E_F - E_C = \frac{\hbar^2}{2m_0} (3\pi^2 n)^{2/3} \quad (2.7.3)$$

The expression is applicable to metals such as copper, gold, etc. In Table 2.2 we show the conduction band electron densities for several metals. The quantity  $E_F$ , which is the highest occupied energy state at 0 K, is called the Fermi energy. We can define a corresponding wavevector

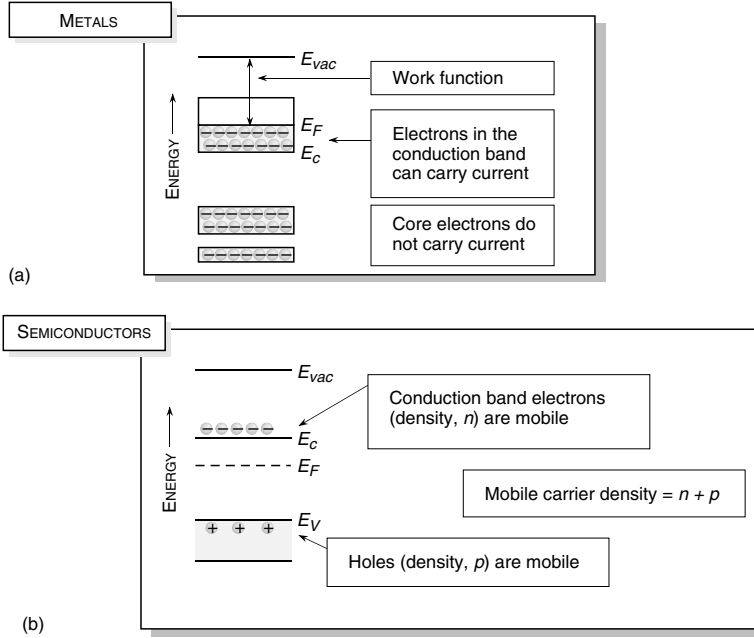


Figure 2.16: (a) In metals the highest occupied band is partially filled and electrons can carry current. (b) A schematic showing the valence band and conduction band in a typical semiconductor. Only electrons in the conduction band and holes in the valence band can carry current.

$k_F$ , called the Fermi vector, and a velocity  $v_F$ , called the Fermi velocity as

$$\begin{aligned} k_F &= (3\pi^2 n)^{1/3} \\ v_F &= \left( \frac{\hbar}{m_0} \right) (3\pi^2 n)^{1/3} \end{aligned} \quad (2.7.4)$$

It is important to note that even at 0 K, the velocity of the highest occupied state is  $v_F$  and not zero, as would be the case if we used classical statistics. At finite temperatures, the Fermi level is approximately given by

$$E_F(T) = E_F(0) \left[ 1 - \frac{\pi^2}{12} \frac{(k_B T)^2}{(E_F(0))^2} \right] \quad (2.7.5)$$

where  $E_F(T)$  and  $E_F(0)$  are the Fermi levels at temperatures  $T$  and 0 K, respectively. In Metals there is very little change in the Fermi level with temperature.

| ELEMENT | VALENCE | DENSITY<br>(gm/cm <sup>3</sup> ) | CONDUCTION<br>ELECTRON DENSITY<br>(10 <sup>22</sup> cm <sup>-3</sup> ) |
|---------|---------|----------------------------------|------------------------------------------------------------------------|
| Al      | 3       | 2.7                              | 18.1                                                                   |
| Ag      | 1       | 10.5                             | 5.86                                                                   |
| Au      | 1       | 19.3                             | 5.90                                                                   |
| Na      | 1       | 0.97                             | 2.65                                                                   |
| Fe      | 2       | 7.86                             | 17.0                                                                   |
| Zn      | 2       | 7.14                             | 13.2                                                                   |
| Mg      | 2       | 1.74                             | 8.61                                                                   |
| Ca      | 2       | 1.54                             | 4.61                                                                   |
| Cu      | 1       | 8.96                             | 8.47                                                                   |
| Cs      | 1       | 1.9                              | 0.91                                                                   |
| Sn      | 4       | 7.3                              | 14.8                                                                   |

Table 2.2: Properties of some metals. In the case of elements that display several values of chemical valence, one of the values has been chosen arbitrarily.

**EXAMPLE 2.1** A particular metal has  $10^{22}$  electrons per cubic centimeter. Calculate the Fermi energy and the Fermi velocity (at 0 K).

The Fermi energy is the highest occupied energy state at 0 K and is given by (measured from the conduction bandedge)

$$\begin{aligned} E_F &= \frac{\hbar^2}{2m_0} (3\pi^2 n)^{2/3} \\ &= \frac{(1.05 \times 10^{-34})^2 [3\pi^2 (10^{28})]^{2/3}}{2(0.91 \times 10^{-30})} = 2.75 \times 10^{-19} \text{ J} \\ &= 1.72 \text{ eV} \end{aligned}$$

The Fermi velocity is

$$\begin{aligned} v_F &= \frac{\hbar}{m_0} (3\pi^2 n)^{1/3} \\ &= \frac{(1.05 \times 10^{-34} \text{ J.s})(3\pi^2 \times 10^{28} \text{ m}^{-3})^{1/3}}{0.91 \times 10^{-30} \text{ kg}} = 7.52 \times 10^5 \text{ m/s} \\ &= 7.52 \times 10^7 \text{ cm/s} \end{aligned}$$

Thus, the highest energy electron has a large energy and is moving with a very large speed.

## 2.7.2 Electrons and holes in semiconductors

In pure semiconductors there are no mobile carriers at zero temperature. As temperature is raised, electrons from the valence band are thermally excited into the conduction band, and in equilibrium there is an electron density  $n$  and an equal hole density  $p$ , as shown in figure 2.17a. Note that the density of allowed states has the form

$$N(E) = \frac{\sqrt{2} (m_{dos}^*)^{3/2} (E - E_c)^{1/2}}{\pi^2 \hbar^3} \quad (2.7.6)$$

where  $m_{dos}^*$  is the density of states mass and  $E_c$  is the conduction bandedge. A similar expression exists for the valence band except the energy term is replaced by  $(E_v - E)^{1/2}$  and the density of states exist below the valence bandedge  $E_v$ . Figure 2.17 shows a schematic view of the density of states.

It is important to note that the density of states mass has a special term in indirect gap materials. In direct gap semiconductors  $m_{dos}^*$  is just the effective mass for the conduction band. In indirect gap materials it is given by (see Appendix C)

$$m_{dos}^* = (m_1^* m_2^* m_3^*)^{1/3}$$

where  $m_1^* m_2^* m_3^*$  are the effective masses along the three principle axes. For Si counting the six degenerate  $X$ -valleys we have

$$m_{dos}^* = 6^{2/3} (m_\ell m_t^2)^{1/3}$$

## Chapter 8

# FIELD EFFECT TRANSISTORS

### 8.1 INTRODUCTION

In this and the next chapter we will examine the field effect transistor (FET) and Metal-Oxide-Semiconductor FETs (MOSFETs). These simple devices are majority carrier devices which are relatively simple to fabricate and are extremely versatile. FETs are now made from a wide variety of materials (Si, SiGe, GaAs, InGaAs, GaN, SiC, etc.). Figure 8.1 shows a GaAs-based Metal Semiconductor FET or MESFET.

The basic concept behind the FET is quite simple and is illustrated in figure 8.2. The device consists of an active channel through which electrons (or holes) flow from the source to the drain. The source and drain contacts are ohmic contacts. The conductivity of the channel is modulated by a potential applied to the gate. This results in the modulation of the charge density flowing in the channel. It is important to isolate the gate from the channel so that no current flows into the gate. The gate isolation is done in a variety of ways, leading to a number of different devices. In the MOSFET, the gate is isolated from the channel by an oxide. This is the basis of the silicon devices. In the metal-semiconductor FET or MESFET, the gate forms a Schottky barrier with the semiconductor and the gate current is small in the useful range of gate voltages. In the junction FET or JFET, a  $p$ - $n$  junction is used in reverse bias to isolate the gate. Heterojunction field effect transistors (HFETs) or modulation doped FETs (MODFETs) use a large bandgap semiconductor to isolate the gate from the active channel. In this chapter we will examine the MESFET or JFET devices. In the next chapter we examine MOSFETs.

### 8.2 JFET AND MESFET: CHARGE CONTROL

The operation of JFETs and MESFETs are similar. The key difference is that in a JFET a  $p^+ - n$  structure is used to create a barrier between the metal gate and the semiconductor while in a MESFET the Schottky barrier height is used. We have seen from chapter 5 that the reverse current in a Schottky junction is much larger than that in a  $p$ - $n$  junction. As a result the JFET is especially used in materials for which it is difficult to produce a large Schottky barrier.

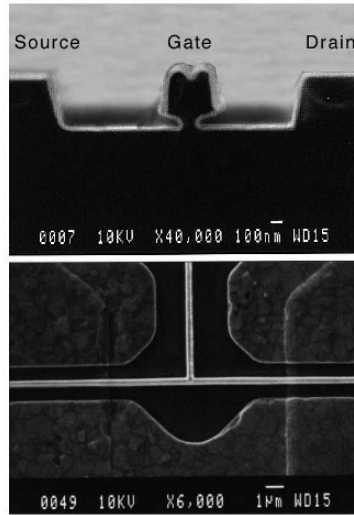


Figure 8.1: MESFETs and JFETs are important devices for high-speed, low-noise amplifiers, D/A and A/D converters, and much “front-end” processing where high speed is critical. These devices exploit materials, like GaAs, InP, and InGaAs, that have transport properties that are superior to Si. (Top) A cutoff cross-section of a  $0.1\ \mu\text{m}$  MESFET. (Bottom) Top view of the MESFET.

Let us examine a typical JFET or MESFET structure as shown in figure 8.3a. The device is based on a low-conductivity substrate on which an  $n$ -type region is grown to form an “active” conducting channel of thickness  $h$ . The gate is formed by a  $p^+$  region ( $n^+$  region for a  $p$ -type FET) or a Schottky barrier. The source and drain are ohmic contacts. In figure 8.3b we show a case where the active channel is partially depleted (say at zero gate bias). A negative bias on the gate reverse biases the gate-ohmic conductor junction (for an  $n$ -type device) and alters the width of the depletion region. This allows the gate to modulate the conductance of the device. In figure 8.3c we show a case where the channel is completely depleted.

Consider the cases shown in figure 8.4. In figure 8.4a we show the device with a small source-drain bias  $V_{DS}$  and no gate bias. As the gate bias is increased and the gate semiconductor junction is reverse biased, the current through the channel decreases until eventually the channel is “pinched off” and there are no free carriers in it. If the gate bias is fixed and the drain bias is increased, as shown in figure 8.5, the gate semiconductor junction near the drain becomes more reverse biased. Eventually, the channel is pinched off near the drain side. At this point the current cannot increase even if the drain voltage is increased. This is called the saturation region. Once the device reaches saturation, the current in the channel remains more or less unchanged.

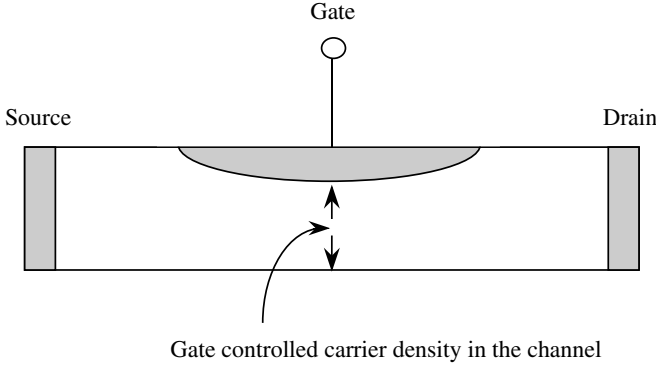


Figure 8.2: The physical principle behind the FET involves the use of a gate to alter the charge in a channel. Depending upon the method used for isolation of gate different FETs arise.

The thickness of the doped active channel is  $h$  as shown in figure 8.3 and  $V_{bi}$  is the built-in voltage. The gate bias required to pinch off the channel is simply given by the depletion approximation

$$h = \left\{ \frac{2\epsilon(V_{bi} - V_{GS})}{eN_d} \right\}^{1/2} \quad (8.2.1)$$

It is possible that the built-in voltage may by itself pinch the channel off. In an  $n$ -channel device, if the device is not pinched off by  $V_{bi}$ , then a negative gate bias will cause pinch-off. In a  $p$ -channel device, a positive bias is needed for pinch-off.

The pinch-off voltage  $V_p$  (called the intrinsic pinch-off voltage) is defined by

$$V_p = \frac{eN_d h^2}{2\epsilon} \quad (8.2.2)$$

and the gate bias needed for pinch-off for the  $n$ -channel device is

$$V_T = V_{bi} - V_p \quad (8.2.3)$$

where  $V_T$  is called the threshold voltage for the device. If the voltage  $V_p$  is smaller than the built-in potential  $V_{bi}$ , the device channel is completely depleted in the absence of a gate bias. A positive gate bias (for  $n$ -channel devices) can allow the channel to have free charge and be conducting. Such devices are said to be enhancement-mode devices. On the other hand, if  $V_p$  is larger than  $V_{bi}$ , the device has free charge in the channel at  $V_G = 0$  since the channel is only partially depleted. A negative gate bias will then turn the device off, i.e., deplete the channel. Such devices are said to operate in the depletion mode. Electronic circuits may use enhancement- or depletion-mode devices or even combinations of them, depending upon the application.

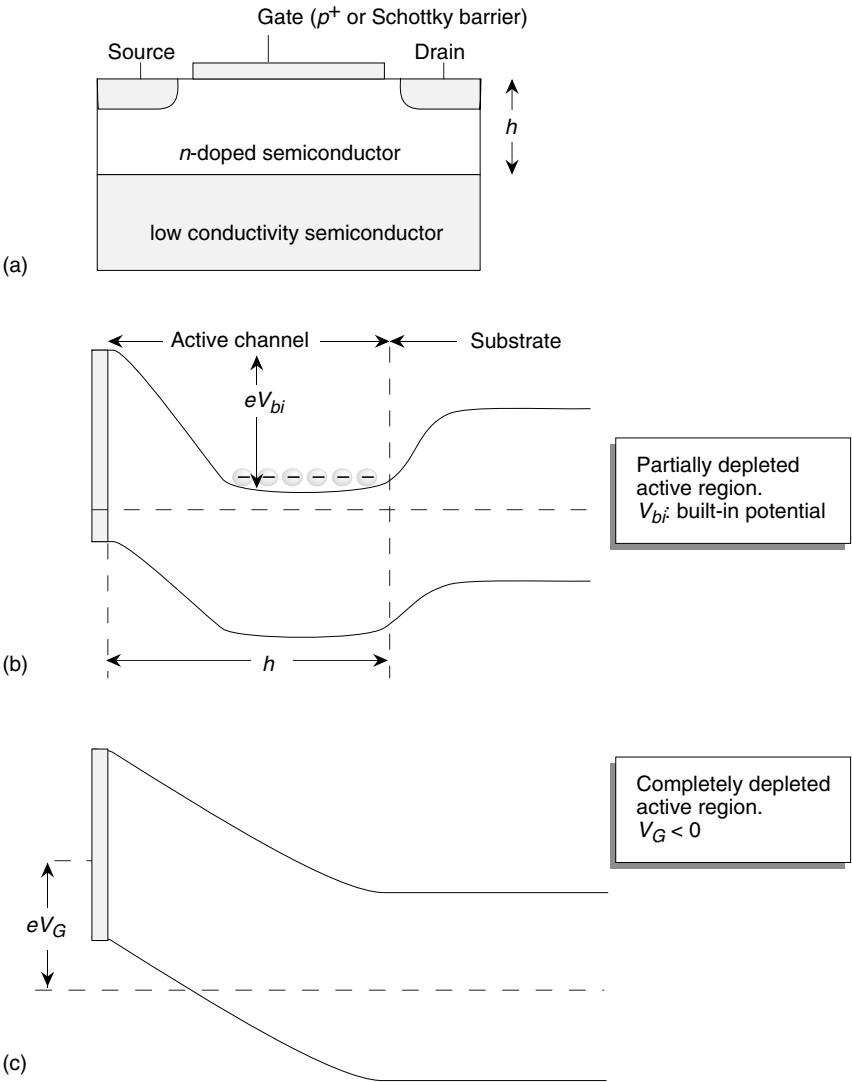


Figure 8.3: (a) A schematic of a JFET or MESFET showing the source, drain, and gate. The channel width of this device is  $h$ ; (b) the band profile when the applied gate bias is zero as is the source-drain bias; (c) the band profile with a negative gate bias so that the channel is depleted.

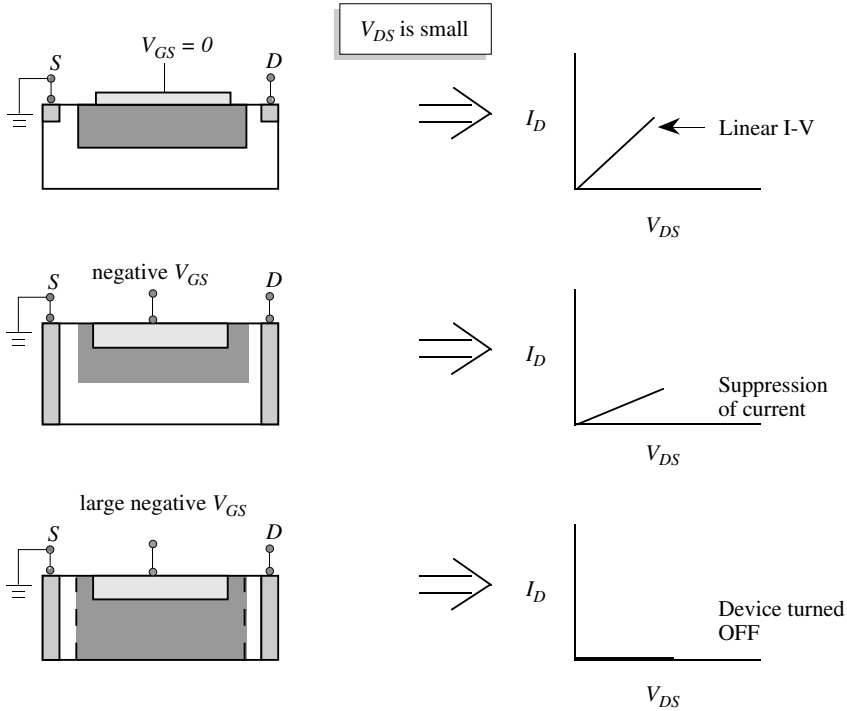


Figure 8.4: (a) Depletion width and channel in a JFET or MESFET under zero gate bias. The channel has a large opening. Such a device is called a depletion-mode device; (b) the device with a negative gate bias showing reduction in the channel opening and current; (c) the gate bias is large and negative and the channel is pinched off with current in the channel zero.

An important consideration in JFET or MESFET technologies is that the gate current be negligible. This requires that the gate be biased appropriately. This also requires a large built-in voltage or Schottky barrier height. For small gap semiconductors (e.g. InGaAs, InSb, etc.) this may not be possible.

**Example 8.1** Consider an  $n$ -MESFET made from GaAs doped at  $10^{17} \text{ cm}^{-3}$ . Calculate the gate current density under normal operation if:

- the gate is made from a Schottky metal with a barrier  $\phi_b = 0.8 \text{ V}$ ;
- the gate is made from a heavily doped  $p^+$  GaAs.

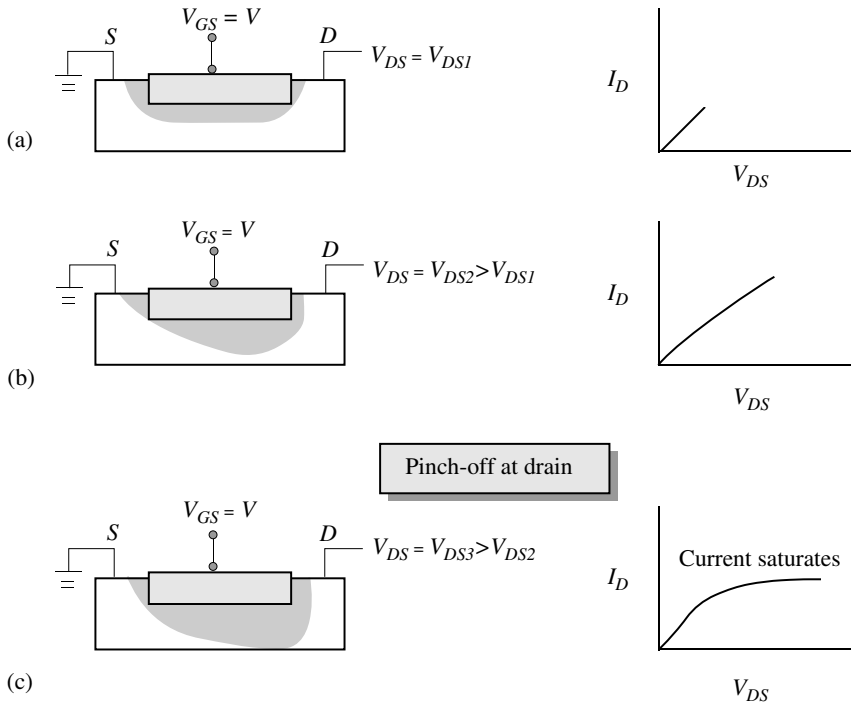


Figure 8.5: The effect of increased drain bias at a fixed gate bias. (a) the drain bias is small; (b) the drain bias is increased and the channel is constricted near the drain; (c) the drain bias is increased to the point that the channel is pinched off at the drain side. The drain current saturates as shown.

You may use the following parameters:

$$\begin{aligned} D_p &= 20 \text{ cm}^2/\text{s} \\ L_p &= 1.0 \text{ } \mu\text{m} \\ A^* &= 8 \text{ A cm}^{-2} \text{ K}^{-2} \end{aligned}$$

The gate current under normal operation is just the reverse-bias current of the junction between the gate and the semiconductor. For the Schottky case we have (see chapter 5)

$$J_s = A^* T^2 \exp\left(-\frac{e\phi_b}{k_B T}\right)$$

This gives

$$J_s = 8 \times (300)^2 \times 4.34 \times 10^{-14} = 3.125 \times 10^{-8} \text{ A/cm}^2$$

For the  $p^+$ -gate we have from  $p$ - $n$  diode theory (see chapter 4)

$$J_0 = \frac{eD_p p_n}{L_p}$$

This gives

$$J_0 = \frac{1.6 \times 10^{-19} \times 20 \times 3.38 \times 10^{-5}}{10^{-4}} = 1.08 \times 10^{-18} \text{ A/cm}^2$$

We see that the gate current is much smaller for the JFET case. However for the GaAs case considered here, the MESFET gate current is small enough for most applications.

### 8.3 CURRENT-VOLTAGE CHARACTERISTICS

The MESFET is one of the simplest three terminal devices to fabricate and to conceptually understand. The most common material used in MESFETs is GaAs. Other compound semiconductors can also be used although it is common to use a HFET approach for most materials. It is important that one have a high resistivity substrate to avoid current flow through the substrate. This is usually done by impurity doping. These impurities create levels at midgap, pinning the Fermi level..

We will first present a very simple model for the current-voltage relation in the MESFET. Then we will describe a more accurate model. However, to obtain realistic results one needs to use computer simulation tools.

In figure 8.6 we show the device structure along with the band profile under the gate. In figure 8.7 we show the MESFET cross-section along with the depletion width under the gate region. In the absence of any bias, there is a uniform depletion region under the gate region, as shown in figure 8.7a. If the gate bias is made more negative, the depletion width spreads further into the active region until eventually the channel is completely depleted. Thus, as the gate bias is increased (to negative values), the total charge available for conduction decreases until the channel is pinched off.

If the gate bias is fixed and the drain voltage is increased toward positive values, current starts to flow in the channel. The depletion region now becomes larger near the drain side, as shown in figure 8.7b. As the drain voltage is increased, the depletion width toward the drain end starts to increase, since the potential difference between the gate and the drain end of the channel increases. The channel then starts to pinch off at the drain end. As this happens, the current starts to saturate. Once the drain voltage reaches a value  $V_{DS}(sat)$  such that the channel pinches off at the drain end, the current remains essentially constant even as the drain voltage is increased.

#### 8.3.1 The Ohmic Regime

As noted earlier we will start with a very simple model which, though not accurate for modern short gate devices, provides insight into the device operation. We will use the following

and

$$V(L_{II}, h) = V_D \quad (8.3.37)$$

The voltage drop in the pinched region (Region II) is  $V_D - (V_G - V_T)$  or  $V_D - V_{DS}(sat) = V_{dp}$ . Therefore

$$L_{II} = \frac{2h}{\pi} \sinh^{-1} \left[ \frac{V_{dp}\pi}{2\mathcal{E}_c h} \right] \quad (8.3.38)$$

and

$$\frac{\partial L_{II}}{\partial V_D} = \frac{1}{\mathcal{E}_c} \left[ 1 + \left( \frac{V_{dp}\pi}{2\mathcal{E}_c h} \right)^2 \right]^{-1/2} \quad (8.3.39)$$

For the case of interest, where the device is biased well into saturation, the second term in the bracket becomes large, giving

$$r_d = \frac{1}{g_d} = \frac{\pi V_{dp}}{2I_D} \left( \frac{L}{h} \right) \quad (8.3.40)$$

This is a very important result. It says that to maintain a high value of  $r_d$ , which is necessary for high voltage gain, it is essential to maintain a high aspect ratio of the gate length to channel thickness ( $L/h$ ). Typically, ( $L/h$ ) should be at least 10. Of course, as the current in the channel decreases,  $r_d$  increases, but this benefit is largely negated by a similar decrease in the device  $g_m$ . An increase in  $r_d$  with  $V_{dp}$  is based on the increase in the saturated region which further isolated the drain potential from the source, reducing  $g_d$ .

## 8.4 HFETs: INTRODUCTION

In the previous sections we have discussed the MESFET (or JFET) devices. In the MESFET the gate is insulated from the channel by a barrier created by either a Schottky barrier (or a  $p^+n$  junction). The charge in the channel is provided by dopants in the channel. The dopants, while providing charge, also cause scattering and reduce mobility. The question arises: Can we have channel charge but avoid dopant scattering? This is possible in the Si MOSFET where charge can be induced by inversion. However, the MOSFET charge has to contend with interface roughness scattering. In the Si/SiO<sub>2</sub> case the interface is between a high quality semiconductor and a non-epitaxial layer and the interface scattering can greatly reduce mobility. Thus while room temperature electron mobility in pure Si is  $\sim 1300 \text{ cm}^2/\text{V.s}$ , it is only half this value in the NMOS channel. It would be ideal to have a heterostructure grown epitaxially where band inversion could occur. However, so far this has been difficult, though advances continue to be made. It is possible to make heterostructure devices where mobile charge is donated by dopants or other fixed charges.

The most widely used heterostructure FET utilizes the modulation doping concept. The device is called modulation doped field effect transistor (MODFET) or high electron mobility transistor (HEMT) or 2-dimensional gate field effect transistor (TEGFET), etc. It has also been shown that polar charges created at interfaces by piezoelectric and/or Spontaneous polarization can also be exploited to create free charge. This approach has become quite dominant in nitride based devices. Heterostructure field effect transistors (HFETs) offer many advantages over MESFETs

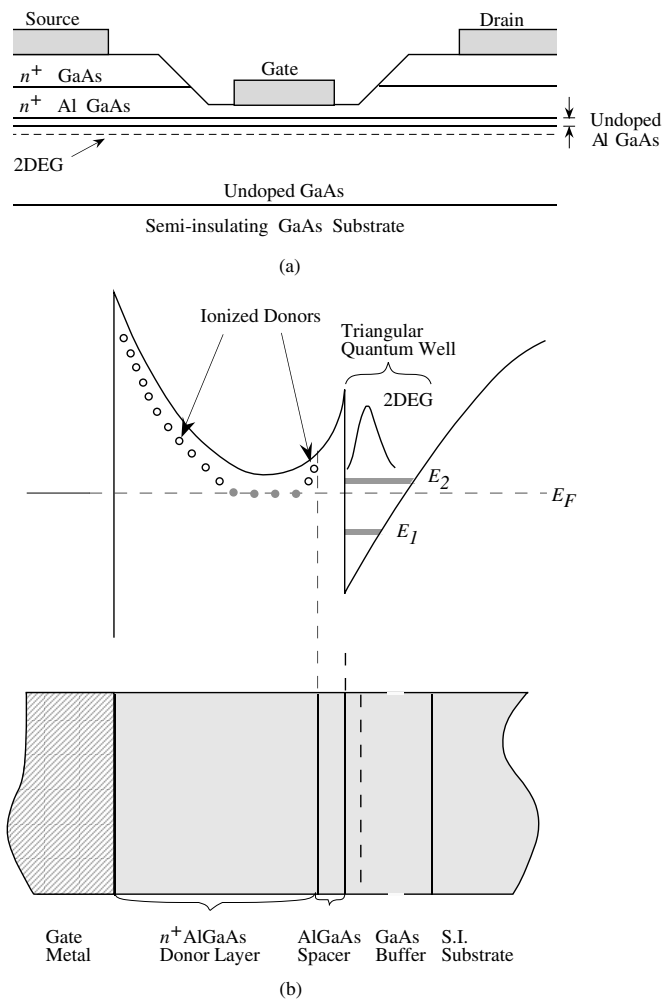


Figure 8.13: (a) A schematic of a GaAs/AlGaAs MODFET. In the figure shown, the gate is “recessed” to have a better control over the 2-dimensional electron gas (2DEG). (b) The band profile of an *n*-MODFET showing bandbending leading to a triangular quantum well at the GaAs/AlGaAs interface.

or MOSFETs. A typical modulation doped device structure is shown in figure 8.13a. We show a structure fabricated by epitaxial techniques such as MBE or MOCVD and using the recessed gate technology. For the AlGaAs/GaAs structure the substrate is semi-insulating GaAs on which an undoped GaAs layer is grown. A heterostructure is formed by depositing AlGaAs which is left undoped to provide a “spacer” region. The remaining barrier material is doped strongly. Finally, a heavily doped GaAs cap layer is deposited on which the ohmic source contacts are deposited. The cap layer is etched off and the Schottky gate is deposited on the high barrier material.

The electrons from the donor atoms in the high barrier material spill over into the low bandgap material conduction band creating a dipole layer. As a result, the band bends as shown in figure 8.13b to produce a quantum well in which the electrons are trapped. The quantum well has a triangular form and the electrons have 2-dimensional properties; i.e., they are free to move in the plane of the device but are confined in the device growth direction. As a result the density of states of the electrons have the usual 2-dimensional features. The term 2-dimensional electron gas (2DEG) is used to describe the electron system.

The key motivations for HFETs are:

- **High Mobility Due to Suppression of Ionized Impurity Scattering:** We have earlier discussed the effect of ionized impurity scattering on mobility. In the HFET, due to the physical separation of the dopants from the free electrons, the mobility is greatly improved. For example, in a GaAs MESFET channel, doped at  $5 \times 10^{17} \text{ cm}^{-3}$ , the room temperature mobility is  $\sim 4000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . In a MODFET channel with equivalent charge density the mobility is essentially limited by phonon scattering to  $\sim 8000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . The effects are even more dramatic at low temperatures.

The improved mobility allows the device to have a very low resistance between the source and the gate region (low access resistance or source resistance). The high field transport in the MODFET channel is, however, not too much better than the MESFET channel since at high fields, transport is governed primarily by phonon (lattice vibration) scattering.

- **Superior Low Temperature Performance:** We had noted in Chapter 3 the carrier freezeout effect that occurs in doped semiconductors at low temperatures. In a MODFET channel, this effect is avoided since the electrons are in a region of energy below the donor levels in the high bandgap material. Thus a high carrier density can be maintained at very low temperature exploiting the low temperature improvement in transport. Extremely low noise, high gain microwave devices are exploiting this low temperature feature for special applications such as deep space signal reception.
- **Use of Superior Materials in the Channel:** In the MODFET, the active channel in which the transport takes place need only be  $\sim 200 \text{ \AA}$ . Thus one can use a very high mobility material system in the channel. Normally materials like InAs or InSb which have very high mobilities cannot be used as MESFETs since it is difficult to process these narrow bandgap materials which are very “soft” and defect prone. However, when only a narrow region is used, the device can be quite robust. On GaAs substrates one can use  $\text{In}_x\text{Ga}_{1-x}\text{As}$  channels for active regions while on InP one can use  $\text{In}_{0.53+x}\text{Ga}_{0.47-x}\text{As}$  as active channel materials. In the two cases above, if  $x \neq 0$ , the channel is under strain, resulting in pseudomorphic MODFETs.

- **High Sheet Charge Density:** The charge density in the 2-dimensional HFET channel depends upon the doping density in the large bandgap material or on polar charge at the interface and the conduction band discontinuity at the channel-barrier interface. By using materials with large conduction band discontinuities, a very high sheet charge density ( $\gtrsim 10^{13} \text{ cm}^{-2}$ ) can be introduced. This results in a very large device transconductance and device performance.

In figure 8.14, we show an SEM image of a state-of-the-art InP-based HFET along with its layer structure,  $I$ - $V$  characteristics, and an integrated circuit composed of these devices. A careful examination of the gate in the SEM image shows that the gate is recessed; the advantages of this are described later in this chapter. The T-gate structure, which is characteristic of all modern high speed FETs, is desirable because it is possible to achieve a very small intrinsic gate length (in this case  $0.1 \mu\text{m}$ ) while still maintaining a bulkier gate metal, which reduces the lateral gate resistance. Also evident is the dielectric passivation layer which covers the device. This prevents undesirable charging of surface states as well as protects the device from contaminants that may be present in the ambient environment.

In the  $I$ - $V$  characteristics, we see that the current saturates at a very low voltage, indicative of the low contact resistance, access resistance, and channel resistance that can be achieved with this technology. However, one can see that the current does not completely saturate. This non-zero output conductance results from short channel effects. Additionally, in the  $I$ - $V$  curves, the device is only biased to 1.5 V, since the breakdown voltage for InP devices with such short gate lengths is typically  $\sim 3$  V. In GaN-based HFET technology, much higher breakdown voltages can be achieved due to the wide bandgaps of the materials in the nitride system.

In this chapter we will examine some important issues in HFETs. In particular we will examine how polar charge can be exploited to create free electron or hole gas. Such undoped HFETs have become very important due their use in the large bandgap AlGaIn/GaN technology.

## 8.5 CHARGE CONTROL MODEL FOR THE MODFET

In a MODFET, electrons are introduced into the channel via doping of a region which is spatially separated from the channel, as shown in figure 8.13 and figure 8.15. In this way, the electron mobility is not degraded by ionized impurity scattering. A number of different doping schemes are possible for this device. The entire barrier material, with the exception of a thin spacer layer near the channel, can be doped, resulting in the structure that was shown in figure 8.13. Alternatively, one can dope a very thin ( $\sim 10 \text{ \AA}$ ) layer of barrier material separated from the channel by an undoped spacer layer. This scheme, known as  $\delta$ -doping, is illustrated in figure 8.15a.

While the continuous doping scheme is in practice easier to implement,  $\delta$ -doping is preferable because the maximum amount of charge that can be induced in the channel is higher. Additionally,  $\delta$ -doping reduces the risk of inducing a parasitic channel within the barrier material. For the MODFET charge control model introduced in this chapter, we assume a  $\delta$ -doped layer with an areal donor density  $N_d [\text{cm}^{-2}]$  separated from the heterointerface by a distance  $d_s$ , as shown in figure 8.15a.

The device is thus a depletion mode MODFET. The 2DEG carrier concentration is given by

$$\begin{aligned} n_s(V_G = 0) &= \frac{12.2 (8.85 \times 10^{-14} \text{ F/cm}) (0.86 \text{ V})}{(1.6 \times 10^{-19} \text{ C}) (350 \times 10^{-8} \text{ cm})} = 1.66 \times 10^{12} \text{ cm}^{-2} \\ n_s(V_G = -0.5) &= \frac{12.2 (8.85 \times 10^{-14}) (0.36)}{(1.6 \times 10^{-19}) (350 \times 10^{-8})} = 6.94 \times 10^{11} \text{ cm}^{-2} \end{aligned}$$

## 8.6 POLAR MATERIALS AND STRUCTURES

### 8.6.1 Polar Materials

An emerging class of materials is the (Al,Ga,In)N-based system for use in both optoelectronics and electronics. These materials are fundamentally different from conventional cubic semiconductors in that they exist normally in the wurtzite phase and exhibit strong polarization in the  $\langle 0001 \rangle$  direction (also known as the  $C$ -direction). Before studying HFETs fabricated from these materials, it is necessary to first understand the effects that these polarization fields have on the electronic properties of the material.

Figure 8.18a shows the ball and stick model of GaN in the  $C^+$  orientation (Ga face on top) and the associated polarization in the crystal. In the classical model, these polarization charges exist on each unit cell. The sum of the internal polarization within the crystal is zero, as shown in figure 8.18b, leaving  $\pm Q_\pi$  charge at each end of the crystal forming a dipole. Since an unscreened dipole will result in a non-sustainable dipole moment, nature will always provide for a screening dipole by placing equal and opposite charges at or close to the charges of polarization dipole, as shown in figure 8.18c. Let us consider some numbers to see how large the polarization dipole moment is.

The spontaneous polarization charge density in GaN  $n_\pi \sim 10^{13} \text{ cm}^{-2}$ . This leads to an electric field

$$\mathcal{E}_\pi = \frac{Q_\pi}{\epsilon} = \frac{en_\pi}{\epsilon} \simeq 1.6 \text{ MV/cm} \quad (8.6.1)$$

In a crystal of thickness  $d = 1 \mu\text{m}$ , the voltage across the material that results from this dipole charge is

$$V_\pi = \mathcal{E}_\pi \cdot d = 160 \text{ V} \quad (8.6.2)$$

which is not sustainable. Hence a screening dipole is essential. This raises the question of what is the nature of the charges that form the screening dipole. They could arise from counter ions from the atmosphere (such as  $\text{H}^+$  and  $\text{OH}^-$ ). This is probably the case for bulk polar materials used in the ceramic industry (such as ZnO for varistors and piezoelectric sensors). However, this is probably not the case for epitaxial GaN thin films, since these films can be created in an atmosphere free of counter ions, such as in an MBE reactor. This begs the question of whether screening is possible without external counter ions. The following discussion addresses this issue.

Consider a lightly doped  $n$ -type GaN sample in the initial stages of growth, shown in figure 8.19a. Due to the lack of availability of GaN substrates, currently GaN is typically grown

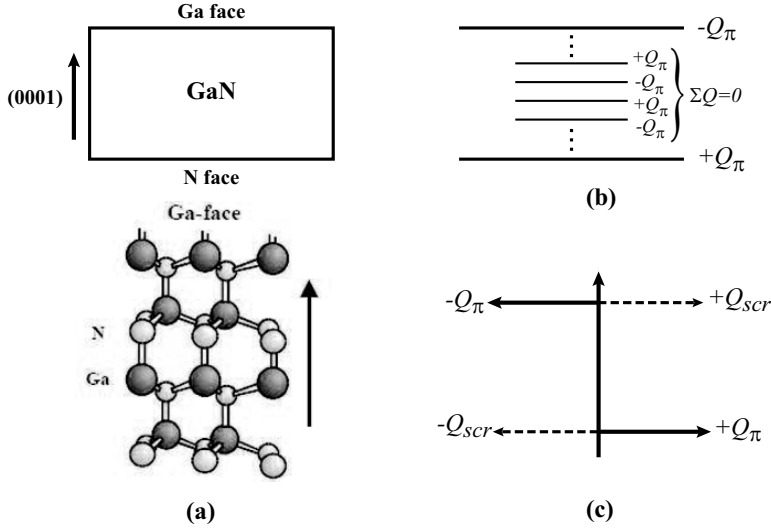


Figure 8.18: (a) Stick-ball representation of Wurtzite GaN crystal structure. (b) Classical model of polarization charge in a polar material such as GaN. (c) Crystal will draw in charge to screen the polarization dipole - From M. J. Murphy et al, MRS Internet J. Nitride Semicond. Res. 4S1, G8.4(1999)

heteroepitaxially (on sapphire, Si, or SiC substrates). The material at the substrate / thin film interface is highly defective and therefore capable of trapping mobile charges. We will assume that the effect of the background  $n$ -type doping on the electric field profile within the material is negligible compared to the electric field generated by the polarization charges. We will also ignore the effects of surface states on the electrical properties of the material. Both of these effects will be considered later.

In the absence of surface states, as the material becomes thicker, the electric field in the material (given by the slope of the conduction and valence band) will remain constant until the valence band crosses the Fermi level, as shown in figure 8.19b. The thickness of the film  $d_{cr}$  at which this occurs is given simply by

$$d_{cr} = \frac{E_g}{e\mathcal{E}_\pi} = \frac{3.4 \text{ eV}}{1.6 \text{ MeV/cm}} \simeq 215 \text{ \AA} \quad (8.6.3)$$

where  $E_g = 3.4 \text{ eV}$  is the bandgap of GaN. Once  $d > d_{cr}$ , holes begin to accumulate at the surface (created by generation across the gap), leading to an equal electron concentration which drifts to the substrate-epi interface (the GaN N-face), creating a screening dipole. This is illustrated in figure 8.19c. The magnitude of the screening charge  $Q_{scr}$  increases continuously

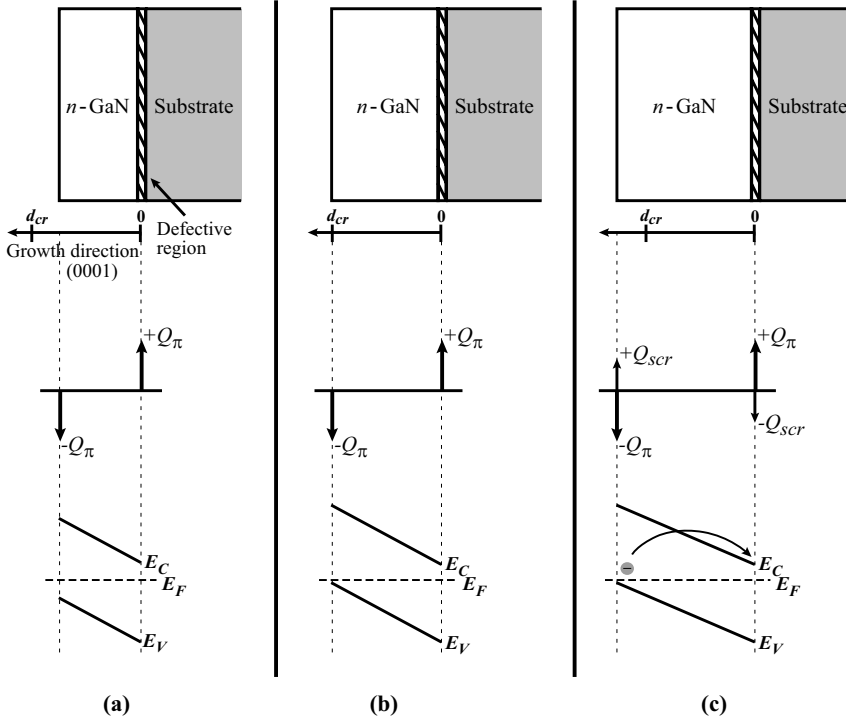


Figure 8.19: Schematic diagram of an *n*-type GaN sample along with charge profile and band diagram (a) during the initial stages of growth, (b) for  $d = d_{cr}$ , and (c) for  $d > d_{cr}$ .

with epitaxial layer thickness. The evolution of the screening charge with distance is obtained by recognizing that the maximum voltage across the structure is the bandgap of the material, or

$$\frac{1}{e} E_g = |\mathcal{E}| \cdot d = \left( \frac{Q_\pi - Q_{scr}}{\epsilon} \right) d \quad (8.6.4)$$

$$Q_{scr} = Q_\pi - \frac{\epsilon E_g}{ed} \quad (8.6.5)$$

As  $d \rightarrow \infty$ ,  $Q_{scr} \rightarrow Q_\pi$ , or in other words for very thick samples the polarization dipole is fully screened.

If we now assume that there exists a surface donor state, a very similar situation develops, except that instead of holes providing the positive screening charge, ionized surface donors do. These states pin the Fermi level at the surface to create a built-in voltage equal to the donor

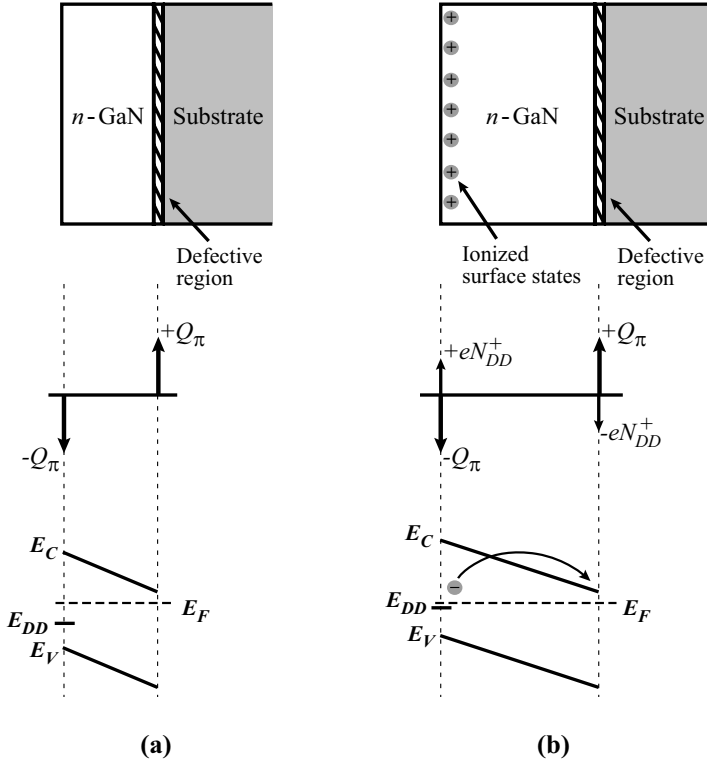


Figure 8.20: Schematic diagram of an *n*-type GaN sample along with charge profile and band diagram when the effects of surface states are taken into account. (a) Very thin GaN, for which surface states are not ionized. (b) Once GaN is thick enough such that  $E_{DD}$  is very close to  $E_F$  at the GaN surface, surface donors become ionized and polarization charge is screened.

depth  $(E_C - E_{DD})/e$ , as illustrated in figure 8.20. As the epitaxial thickness increases, the donor level  $E_{DD}$  approaches the Fermi level  $E_F$  at the GaN surface, and the screening charge  $N_{DD}^+$  increases as given by the Fermi-Dirac occupancy probability

$$N_{DD}^+ = [1 - f(E_{DD}(0))] N_{DD} \quad (8.6.6)$$

$$= \left[ \frac{\exp\left(\frac{E_{DD}(0) - E_F}{k_B T}\right)}{1 + \exp\left(\frac{E_{DD}(0) - E_F}{k_B T}\right)} \right] \cdot N_{DD} \quad (8.6.7)$$

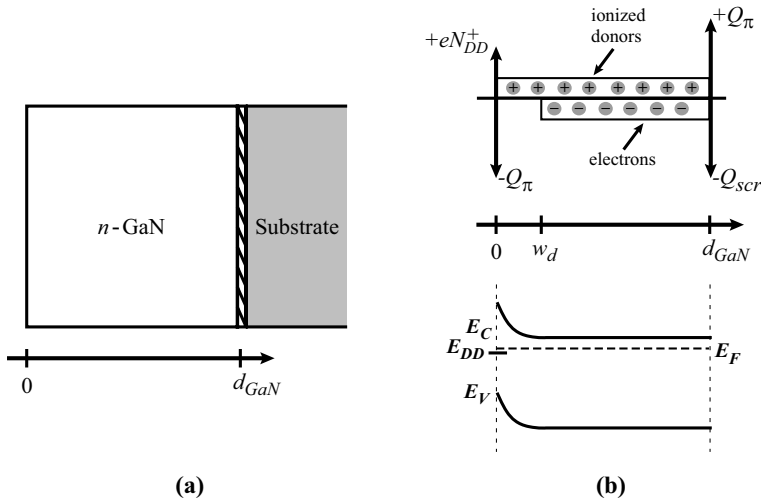


Figure 8.21: Schematic diagram of (a) a thick  $n$ -type GaN sample along with (b) the corresponding charge profile and band diagram when surface states and the donor charge are taken into account.

If  $eN_{DD} > Q_\pi$ , then  $Q_\pi$  will be fully screened when  $(E_C - E_{DD}) - (E_C - E_F)$  is very close to zero, or in other words when  $E_{DD}$  is very close to  $E_F$ . Analogous to the previous case of screening via holes,

$$N_{DD}^+ = Q_\pi - \frac{\epsilon E_{DD}}{ed} \quad (8.6.8)$$

where  $E_{DD}/e$  is now the built-in voltage as opposed to  $E_g/e$ .

We have ignored the effects of the donor charge in the analysis until now because we were seeking to understand the formation of the screening dipole. Figure 8.21 shows the band diagram of a thick  $n$ -type GaN film. Experimental evidence has shown that the surface of GaN indeed has a neutral level,  $E_{DD}$ , which is currently assumed to be the position of the surface donor. Since the GaN is considered to be very thick, the energy bands must be flat (zero electric field). The surface negative polarization charge  $-Q_\pi$  is balanced by the sum of the positively charged ionized surface states  $N_{DD}^+$  and the areal density of charges in the depletion region  $N_d \cdot w$ , or

$$Q_\pi = N_{DD}^+ + N_d \cdot w \quad (8.6.9)$$

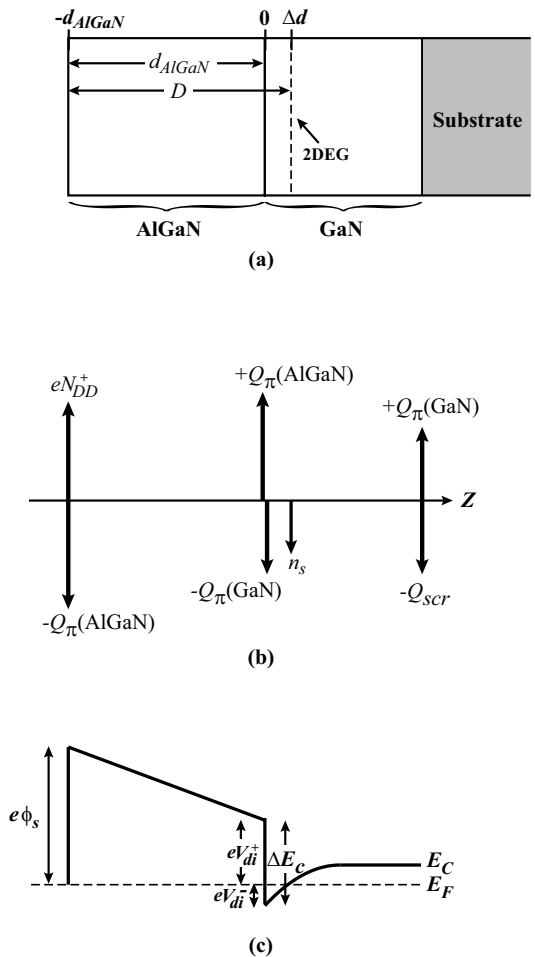


Figure 8.22: Polar heterostructures can generate a 2DEG which is used as the channel region of an HFET. (a) Typical AlGaIn/GaN heterostructure used in polar HFET technology, along with (b) the charge distribution and (c) the band diagram of the structure.

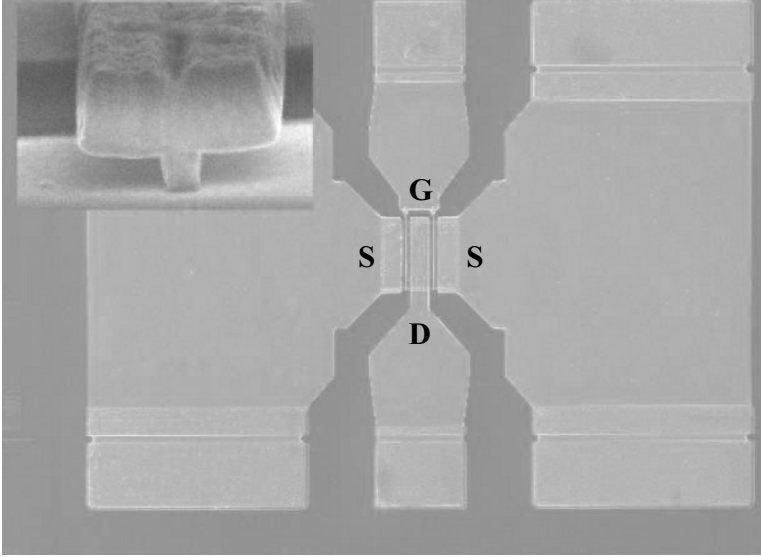


Figure 8.23: Top view of an AlGaIn/GaN HFET structure with 2 gate fingers. Pictured in the inset is a close up of the  $0.12\ \mu\text{m}$  T-gate. Picture courtesy of Ilesanmi Adesida.

### 8.6.2 Polar HFET Structures

Now that we have described how charge is distributed within polar materials, we are ready to show how polarization fields can be used to generate a 2DEG in polar heterostructures. Consider the AlGaIn/GaN structure illustrated in figure 8.22. The charge density distribution is shown along with the band diagram. For sufficiently thick AlGaIn layers, the surface potential  $e\phi_s$  is pinned by the surface donor and is approximately equal to the donor depth  $E_{DD}$ . Due to the lattice mismatch between AlGaIn and GaN, the thin AlGaIn cap is under tensile strain. Hence the total polarization charge at the AlGaIn surface  $-Q_\pi(\text{AlGaIn})$  is the sum of the spontaneous and piezoelectric contributions from the AlGaIn. In addition to the negative polarization charge, there will also be a positive charge at the surface  $N_{DD}^+$  resulting from the ionized surface donors.

At the AlGaIn/GaN interface, the net polarization charge  $Q_\pi(\text{net})$  is the sum of the polarization contributions from the AlGaIn and the GaN, or

$$Q_\pi(\text{net}) = Q_\pi(\text{AlGaIn}) - Q_\pi(\text{GaN}) \quad (8.6.10)$$

$Q_\pi(\text{net})$  is a positive number for Ga-face polarity because of the higher polarization in the AlGaIn relative to GaN. From the band diagram, we can see that there must also exist a distribution of electrons in the GaN near the AlGaIn/GaN heterointerface. Again, we have drawn

the distributed charge as a 2 dimensional sheet charge of density  $en_s$  a distance  $\Delta d$  from the heterointerface, where  $\Delta d$  is the centroid of the charge distribution.

To find our 2DEG carrier density  $n_s$ , we follow the same methodology as in our MODFET analysis (see equation 8.5.5). Doing this, we get

$$\phi_s - V_1 - \frac{\Delta E_c}{e} + V_{di}^- = 0 \quad (8.6.11)$$

where  $V_{di}^-$  was given in equation 8.5.3 and  $V_1$  is given by

$$V_1 = \frac{[Q_\pi(\text{net}) - en_s] d_{AlGaN}}{\epsilon} \quad (8.6.12)$$

Substituting these values into equation 8.6.11 and setting  $d_{AlGaN} + \Delta d = D$  gives us for  $n_s$

$$n_s = \frac{Q_\pi(\text{net}) \cdot d_{AlGaN} - \epsilon(\phi_s - \Delta E_c/e)}{eD} \quad (8.6.13)$$

which is the same expression as that derived for conventional HFETs with  $d_s = 0$  (i.e. with the donor sheet at the heterointerface). This is reassuring, since in the case of AlGaIn/GaN heterostructures, the positive sheet charge that induces the 2DEG is the net polarization charge at the heterointerface. The physical difference between conventional and polar HFETs is simply the origin of the electrons in the 2DEG. In conventional HFETs, the channel electrons are provided by a donor sheet, while in GaN-based HFETs, they come from ionized surface donor states.

In an HFET structure, we place a gate metal on top of the AlGaIn layer and apply a gate voltage to modulate the charge in the 2DEG. The only difference between the HFET charge control analysis and the one presented here for an AlGaIn/GaN heterostructure is that in the HFET, the potential barrier at the metal/AlGaIn interface is given by  $\phi_b - V_G$ , where  $\phi_b$  is the metal-semiconductor barrier height and  $V_G$  is the applied gate voltage. Thus, for the AlGaIn/GaN HFET, the 2DEG sheet charge density as a function of gate voltage can be written as

$$n_s(V_G) = \frac{Q_\pi(\text{net}) \cdot d_{AlGaIn} + \epsilon[V_G - (\phi_b - \Delta E_c/e)]}{eD} \quad (8.6.14)$$

## 8.7 DESIGN ISSUES IN HFETS

In addition to the issue of aspect ratio discussed for MESFETs and JFETs in Chapter 8, there are several other design issues to be considered in HFETs. They are summarized in table 8.1. We will discuss a number of techniques which are employed in modern HFET processes that address these issues.

### 8.7.1 $n^+$ Cap Layers

$n^+$  cap layers are used to reduce the contact resistance as well as the access resistance in the device. The schematic of an AlGaAs/GaAs HFET with an  $n^+$  GaAs cap layer is shown in

Design Issues in HFET Technology

| DESIGN ISSUE                     | NEED     | METHODOLOGY                                                            |
|----------------------------------|----------|------------------------------------------------------------------------|
| 1. Ohmic contact resistance      | Minimize | $n^+$ cap layers, optimized alloying schemes, ion implantation         |
| 2. Channel and access resistance | Minimize | Ion implantation, high 2DEG density-mobility product, $n^+$ cap layers |
| 3. Substrate injection           | Minimize | Quantum well structures, $p$ -type buffers                             |
| 4. Gate leakage                  | Minimize | Junction HFETs, insulated gate structures, gate recess, field plates   |
| 5. Parasitic capacitances        | Minimize | Low K dielectrics, lateral structures preferred                        |
| 6. Breakdown voltage             | Maximize | Gate recess structures, high bandgap materials, field plates           |
| 7. Threshold voltage             | Control  | Etch-stop layers, controlled epitaxial growth                          |

Table 8.1: Overview of technology issues that must be addressed in HFET design.

figure 8.24 along with a band diagram. The access resistance of the HFET is comprised of the sheet resistance of the  $n^+$  cap layer  $R_{n^+}$ , the sheet resistance of the 2DEG  $R_{2DEG}$ , and the interchannel resistance posed by the barrier to electron flow between the cap and the channel  $R_{int}$  (see figure 8.24b). The total resistance can be modeled as a distributed network of all of these components, as shown schematically in figure 8.24a.

Solutions to reducing  $R_{int}$  are to reduce the barrier to electron flow from the  $n^+$  layer to the 2DEG channel, and to reduce the barrier to tunneling. The first is best achieved by increasing the doping in the  $n^+$  layer so that it is very degenerate, causing  $E_F$  to rise above  $E_C$ . The second is achieved by adding  $n^+$  doping in the AlGaAs layer nearest to the surface, which in turn enhances the surface electric field and thereby tunneling. A schematic diagram illustrating the benefits of both these solutions is shown in figure 8.25.

8.7.2 Maximizing 2DEG Conductivity

The 2DEG conductivity in MODFET structures  $\sigma$  is given by

$$\sigma = e\mu_n n_s \tag{8.7.1}$$

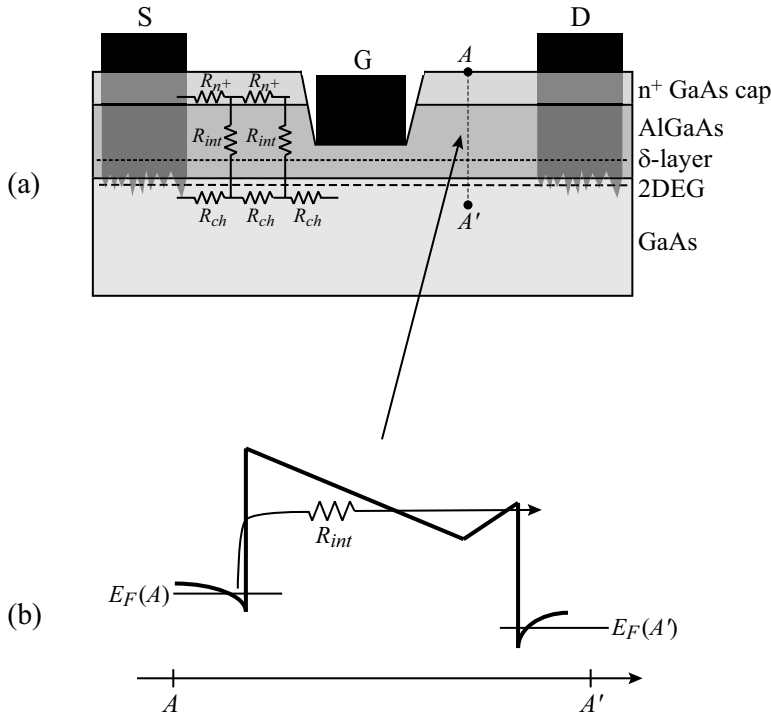


Figure 8.24: (a) Schematic of an AlGaAs/GaAs HFET with an  $n^+$  cap layer and a recessed gate. Also shown are the various resistive components that make up the source access resistance. (b) Band diagram across the structure.

The 2DEG conductivity is at a maximum when the  $\mu_n \cdot n_s$  product is maximized. From our discussion of the Lever Rule in section 8.5, it is clear that the 2DEG density  $n_s$  increases as the  $\delta$ -doping sheet is brought closer to the heterointerface. However, decreasing the spacer distance  $d_s$  also causes the electron mobility  $\mu_n$  to decrease because of the increase in remote ionized impurity scattering. It is therefore clear that the 2DEG conductivity will have a maximum at a value  $d_s(\text{optimum})$  which must be determined for each material system and doping level. Typical values are 5 nm for the AlInAs/GaInAs system, 3 nm for the AlGaAs/InGaAs system, and 2 nm for the AlGaAs/GaAs system.

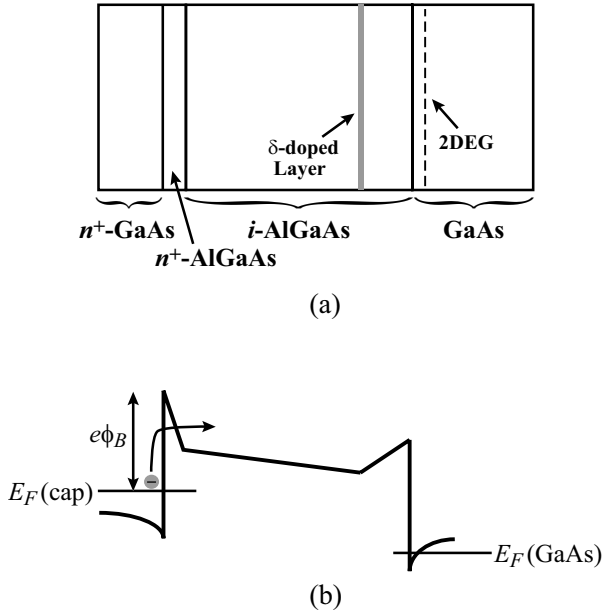


Figure 8.25: (a) Layer structure and (b) band diagram of an AlGaAs/GaAs HFET with a highly degenerate  $n^+$  GaAs cap layer directly above a thin  $n^+$  AlGaAs layer. The high doping in the GaAs cap layer reduces the barrier  $e\phi_B$  that electrons must overcome, and the  $n^+$  AlGaAs layer increases the probability of electrons tunneling through a portion of the barrier.

### 8.7.3 Back-barriers to Substrate Injection

Control of channel charge is the essence of FET operation. If electrons travel through the path labeled  $I_s$  in figure 8.26a, then they are effectively controlled by the gate. Electrons traveling along the path labeled  $I_{par}$  within the substrate and far from the gate are not effectively modulated and are parasitic currents leading to both reduced output resistance (hence low power gain) and low current gain. To keep electrons from being injected into the substrate, we need to present a barrier to current flow, as shown in figure 8.26b. This can be done by introducing a fully depleted  $p$ -type layer or a wider bandgap buffer. The band diagrams for each of these are shown in figure 8.26c and figure 8.26d.

The  $p$ -type buffer introduces negative space charge to the region immediately below the channel, thus increasing the electrostatic barrier to electron injection into the buffer by a maximum amount

$$\Delta\phi_b \simeq \frac{eN_a d_{bar}^2}{2\epsilon} \quad (8.7.2)$$

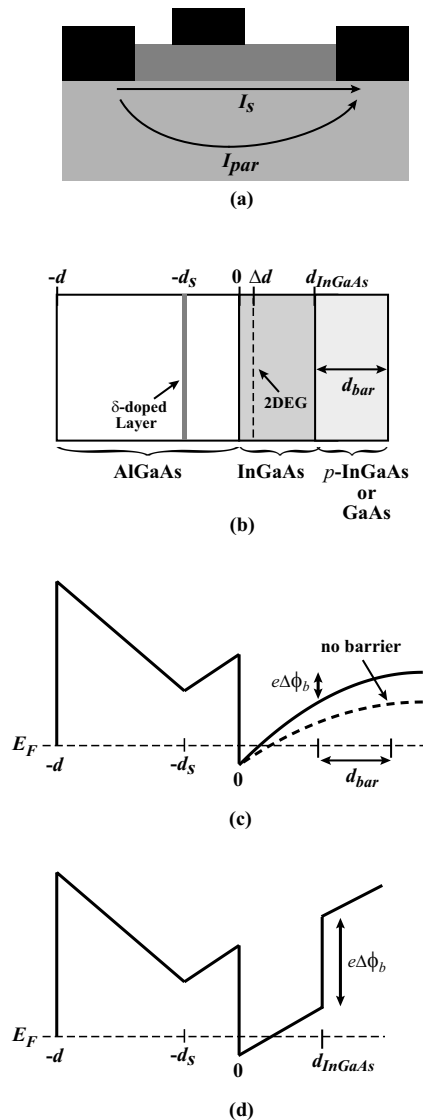


Figure 8.26: (a) In standard MODFET structures, a small parasitic leakage current  $I_{par}$  flows through the substrate. (b) MODFET structure which incorporates a back barrier to prevent current injection through the substrate. (c) Band diagram for a  $p$ -type barrier. (d) Band diagram for a barrier composed of a wide bandgap buffer.

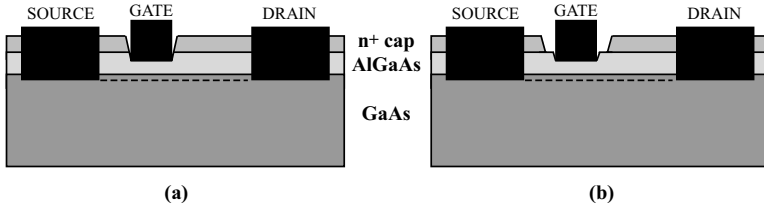


Figure 8.27: HFET structures with (a) a single gate recess and (b) a double recess.

The wide bandgap buffer approach provides a barrier increase of

$$\Delta\phi_b = \frac{\Delta E_C}{e} \quad (8.7.3)$$

The latter confinement scheme is common in AlGaAs/InGaAs/GaAs pseudomorphic HFETs and is the preferred design for GaAs-based HFET structures.

#### 8.7.4 Gate Recess Design

Recessed gate structures are required when  $n^+$  cap layers are employed, and they can also be designed to improve gate leakage and breakdown characteristics as well as to control the device threshold voltage. Designing the gate recess is one of the more important issues in HFET design. Recess structures can generally be placed into two categories: single recess (figure 8.27a) and double recess (figure 8.27b) structures.

The single recess is designed so that the recess and the gate metal are both defined through a single opening in the photoresist such that the recess width is approximately equal to the gate length  $L_g$ . The advantage of this process is that the source and drain access resistances are minimized, so the transit delay is determined dominantly by the gate length, as the high field region in the structure is effectively terminated by the source and drain cap layers. The major disadvantage of this scheme is that the lack of depletion field extension beyond the gate increases the electric field at the drain edge of the gate, thus increasing gate leakage and decreasing the breakdown voltage.

The double recess design, shown in figure 8.27b, allows one to trade off transit delay versus gate leakage and breakdown. By utilizing the first recess of length  $L_R$  to etch through the  $n^+$  cap layer and the second recess to simultaneously define the gate length and threshold voltage.

Single recess structures are used for small signal analog applications such as low noise amplifiers and in digital circuits, whereas double recess designs are used in large signal analog applications such as power amplifiers.

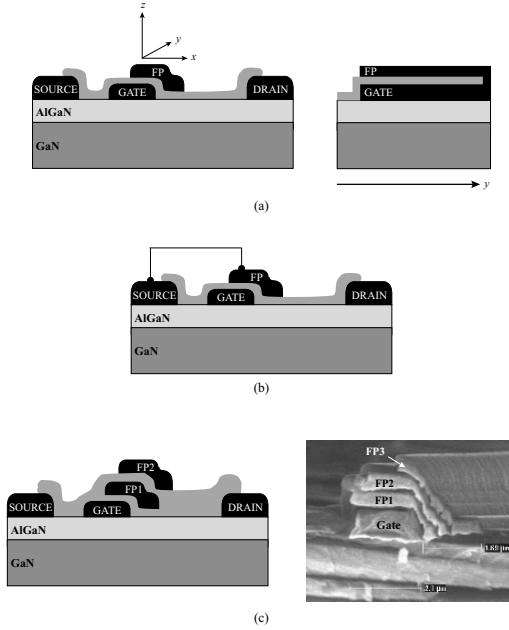


Figure 8.28: Various field plate configurations. (a) Gate-terminated field plate. (b) Source-terminated field plate. (c) Multiple field plate structure. SEM image courtesy of Y. Dora, UCSB.

### 8.7.5 Field Plates

One can actively control the gate extension beyond the drain edge of the gate and thereby reduce the peak electric field by using field plate structures. This is advantageous for applications such as high voltage switching and high power amplifiers, in which very high breakdown voltages are necessary. For this reason, field plates have become especially popular for HFETs in the GaN-based material system. There are a number of methods of implementing field plates, a few of which are shown in figure 8.28. One can have a dielectric-assisted extension of the gate toward the drain (i.e. a gate-terminated field plate). The gate extension effectively modulates the channel beyond the primary gate, thereby spreading the electric field between two peaks, one at the gate edge and the other at the edge of the termination, as shown in figure 8.32. The penalty for this approach is the enhanced gate-drain feedback capacitance  $C_{GD}$ .

Field shaping can also be achieved by utilizing a field plate connected to the source, as shown in figure 8.28b. Here, image charges on the plate result in an enhanced drain-source capacitance

### Comparison of AlGaIn/GaN HEMT to InAlAs/In<sub>0.53</sub>Ga<sub>0.47</sub>As/InP Substrate

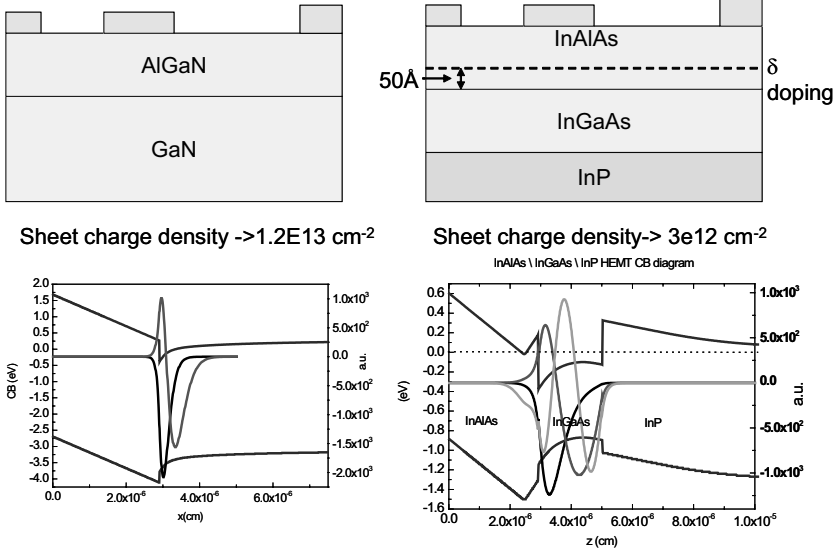


Figure 8.29: Schematic structure and band diagrams of AlInAs/GaInAs and AlGaIn/GaN HFETs.

$C_{DS}$  and an enhanced gate-source capacitance  $C_{GS}$ , but a reduced  $C_{GD}$  because of the screening of the gate from the drain by the source-connected field plate. One can effectively trade off the capacitances based on the geometry of the gate-connected and source-connected field plates, thus mapping out a design space of gain and breakdown voltage.

### 8.7.6 Comparison of two disparate material systems: AlInAs/GaInAs and AlGaIn/GaN

It is instructive to compare the behavior of two families of HFET devices which could in some ways be considered to be at opposite ends of compound semiconductor space. One is the AlInAs/GaInAs/InP system, where the In composition in the GaInAs channel can be increased beyond the 53% value required to achieve lattice matching to come close to the 6.1 Å lattice constant of InAs. The bandgap of course decreases from 0.74 eV in the lattice matched case toward the bandgap of InAs ( $\sim 0.36$  eV). The mobility in the 2DEG can increase from 9,000  $\text{cm}^2/\text{V}\cdot\text{s}$  to over 15,000  $\text{cm}^2/\text{V}\cdot\text{s}$ . The effective mass of the electron decreases from  $0.47m_0$  to  $0.25m_0$ .

At the other extreme is the AlGaIn/GaN system, where  $E_G$  in the channel is 3.4 eV. and the effective mass is  $0.2 m_0$ . Though the channel can be modified in a pseudomorphic fashion by adding In, the advantages are not obvious. Figure 8.29 shows the band structure of the AlGaIn/GaN HFET and the AlInAs/GaInAs HFET. One feature to note is that the AlInAs/GaInAs HFET is modulation doped whereas the AlGaIn/GaN HFET achieves its 2DEG as a result of polarization. The second is that the 2DEG concentration is only  $3 \times 10^{12} \text{ cm}^{-2}$  in the AlInAs/GaInAs HFET, as opposed to  $1.2 \times 10^{13} \text{ cm}^{-2}$  in the AlGaIn/GaN HFET. The reason is that beyond an electron concentration of that order, the conduction band in the AlInAs touches the Fermi level, drastically reducing the modulation efficiency. The low scattering rates in GaInAs because of the small electron effective mass and the large separation between the  $\Gamma$  and  $L$  valleys results in large electron velocity overshoot in channels which are much smaller than the mean free path.

It is imperative to include velocity overshoot in calculating current-voltage ( $I - V$ ) curves of InGaAs HEMTs where an average velocity of over  $4 \times 10^7 \text{ cm/s}$  is easily attained for gate lengths of  $0.1 \mu\text{m}$ . In comparison, the large effective mass of electrons in GaN, the high phonon energy, and the strong coupling between electrons and phonons increases the scattering rate by over an order of magnitude compared to InGaAs ( $10^{13} \text{ s}^{-1}$  vs.  $10^{12} \text{ s}^{-1}$  in bulk materials). Hence, the probability of overshoot is much lower in this case.

Figure 8.30 shows that the GaN HFET has a very small fraction of the  $0.1 \mu\text{m}$  long channel exhibiting velocity overshoot, whereas the InGaAs HFET exhibits it over the full channel. Initial estimates suggest that velocity overshoot will become important at gate lengths of 20 nm or less in the GaN system. The difference in the non-stationary electron transport behavior is the primary reason why the InGaAs HFET shows excellent  $f_T$  behavior with decreasing gate length, as shown in figure 8.30. The current state-of-the-art is an  $f_T$  of over 560 GHz at a gate length of 30 nm. On the other hand, AlGaIn/GaN HFETs have achieved an  $f_T$  value of 163 GHz at 90 nm. The power performance of a state-of-the-art AlGaIn/GaN HFET, which has high breakdown voltage because of the large  $E_g$  is shown in figure 8.31a.

### 8.7.7 Non-idealities in state-of-the-art transistors

The performance of state-of-the-art HEMTs is strongly affected by gate modulation efficiency, electron confinement in the channel and small signal access resistances. This section will show several examples of how these parameters affect the performance of the transistors. It will also describe some techniques that allow a higher performance by overcoming these limitations. We use AlGaIn/GaN HEMTs as the vehicle for demonstration.

As shown in previous sections, a good aspect ratio between the gate length and the gate-to-channel distance is critical to obtain a good modulation of the channel electrons by the gate. This is especially important in high frequency devices where a poor gate modulation efficiency degrades  $f_T$ . To illustrate this problem, figure 8.33 shows the  $f_T$  of AlGaIn/GaN HEMTs for different gate aspect ratios. There is a clear increase in  $f_T$  and  $f_{\text{max}}$  as the aspect ratio increases.

However, a good aspect ratio is not enough to allow a good modulation of the channel electrons by the gate. A poor carrier confinement in the channel can also degrade the performance of transistors, even with good gate aspect ratios. Figure 8.35a shows the transconductance as a

Comparison of predicted  $f_T$  and experimental work

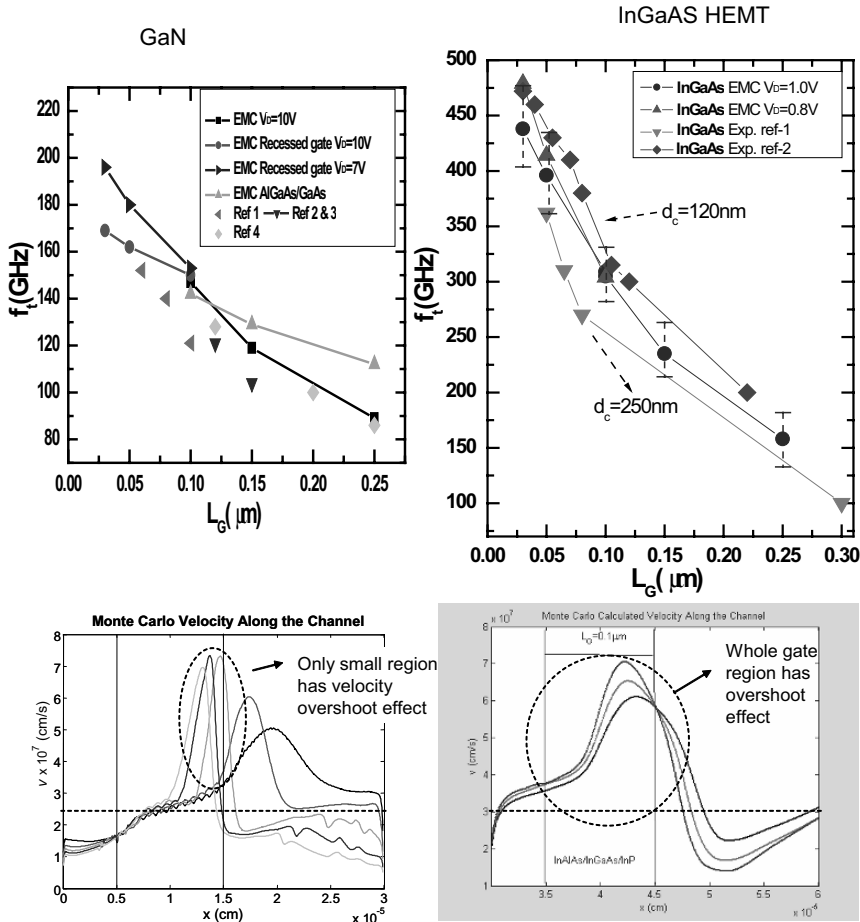


Figure 8.30:  $f_T$  vs.  $L_G$  and velocity field profiles along the channel for both GaN and InGaAs HFETs.

function of gate voltage for different drain voltages in an AlGaIn/GaN HEMT. At low drain voltages, the gate can easily modulate the electrons in the channel and a good pinch-off voltage is obtained for a gate voltage of -5 V. However, as the drain voltage increases, the pinch-off degrades significantly, shifting to lower VGS voltages and becoming softer. These problems are the consequence of the poor electron confinement typical of single heterojunction devices

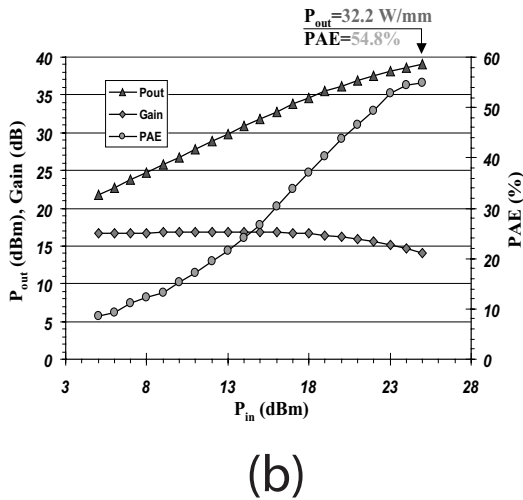
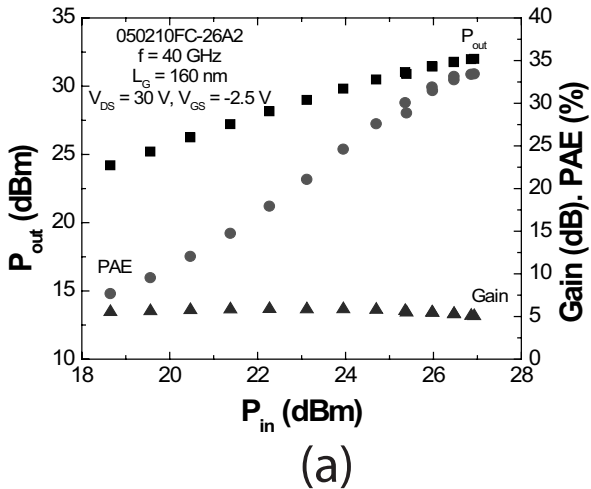


Figure 8.31: (a) Power performance of an AlGaIn/GaN HFET at 40 GHz. The maximum power output of this device  $P_{out} > 10.5$  W/mm with a PAE of 33%. Figure courtesy of T. Palacios, UCSB. (b) Record power densities have been achieved by employing field plates in AlGaIn/GaN technology. Shown here are power measurements taken at 4 GHz of a  $246 \mu\text{m}$  wide device biased at  $V_{DS} = 120$  V. The maximum output power density  $P_{out} = 32.2$  W/mm with a PAE of 54.8%. Figure courtesy of Y.-F. Wu, Cree Inc.

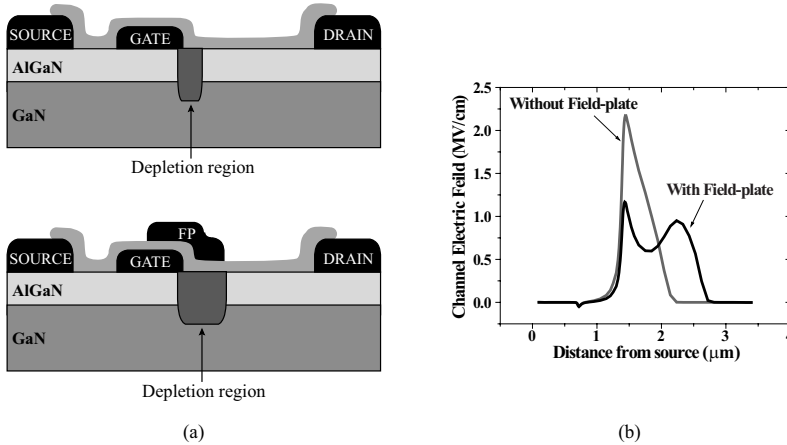


Figure 8.32: (a) Schematic diagrams of HFET structures with and without gate-terminated field plates. In the field plated device, the depletion region extends over a larger lateral distance. (b) Electric field profiles within the depletion region along the channel of both devices.

like the AlGaIn/GaN HEMT. While there is a significant potential barrier at the AlGaIn/GaN heterointerface, there is no barrier between the channel and the buffer. Therefore, it is easy for hot electrons to get injected into the buffer, which increases the gate to channel distance and degrades the performance.

Especially in high frequency devices with very short channels, it is important to increase the confinement of the channel by providing a potential barrier between the channel and the buffer. This barrier can be formed by the conduction band discontinuity between a wide bandgap semiconductor buffer and a narrow bandgap channel. This is the approach normally followed in AlGaAs/GaAs/AlGaAs transistors. The channel confinement can also be increased by doping the buffer  $p$ -type, which generates an electric field in the buffer in a direction that opposes the injection of hot electrons from the channel. An additional option in nitride-based devices is to use ultra thin InGaN backbarrier layers as shown in figure 8.36. In this device, the difference in the polarization coefficients between the GaN buffer and the InGaN backbarrier induces two sheets of fixed charge at the GaN/InGaN interfaces. These polarization induced charges generate an electric field in the InGaN layer which lowers the conduction band in the GaN channel with respect to the GaN buffer. This creates an effective conduction band discontinuity which provides a barrier for the flow of electrons into the buffer as shown in the band diagram in figure 8.36b. The improved confinement provided by the InGaN back barrier allows much better gate modulation at high drain voltages as shown in the transconductance measurements of figure 8.35b. In these improved devices, there is no degradation in the quality of the pinch-off as the drain voltage increases, although there is still a shift in the pinch-off voltage.

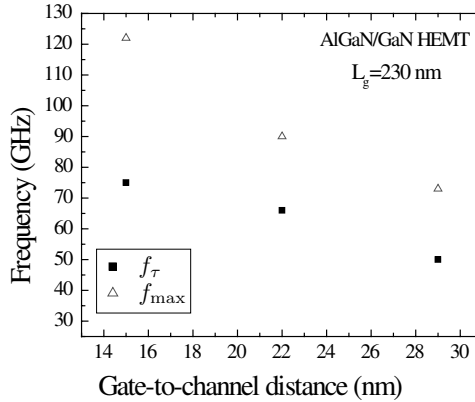


Figure 8.33: Effect of the gate-to-channel distance in the frequency performance of an AlGaIn/GaN HEMT with a gate length of 230 nm.

The higher channel confinement of double heterojunction devices is also beneficial in increasing the output resistance of the transistors. Figure 8.34 compares the output resistance as a function of gate length for several standard AlGaIn/GaN HEMTs and some AlGaIn/GaN HEMTs with InGaIn backbarrier. Almost a 50% increase in the output resistance can be measured in the devices with higher channel confinement. This increase in output resistance also causes an increase in the  $f_{max}$  of the devices, as predicted by equation 8.8.12, and shown in figure 8.37.

Another interesting non-ideality in the behavior of many transistors is the decrease of  $g_m$  with drain current. From equation 8.8.2, the transconductance of a HEMT operating in the saturated mode should be independent of the drain current level. However, this is normally not the case. As shown in figure 8.38 for an AlGaIn/GaN HEMT,  $g_m$  decreases as current increases once that the maximum  $g_m$  has been reached. This kind of behavior has been observed in many different transistor technologies, including Si MOSFETs, AlGaAs/GaAs MODFETs and AlGaIn/GaN HEMTs. The cause of this decrease is different in each technology. In Si MOSFETs, simulations have related this decrease in performance with roughness at the Si/SiO<sub>2</sub> interface. On the other hand, in AlGaAs/GaAs HEMTs, as the drain current increases, there is a reduction in the modulation efficiency of the gate due to the capture of channel electrons by the ionized donors in the AlGaAs barrier, which reduced  $g_m$ . Finally, in GaN technology, the reason for this decrease is related to the increase in small signal source access resistance due to a reduction in the electron mobility at higher electric fields in this material system. Other studies have also proposed the emission of hot phonons and the subsequent reduction of the electron velocity as a possible cause for this decrease in performance.

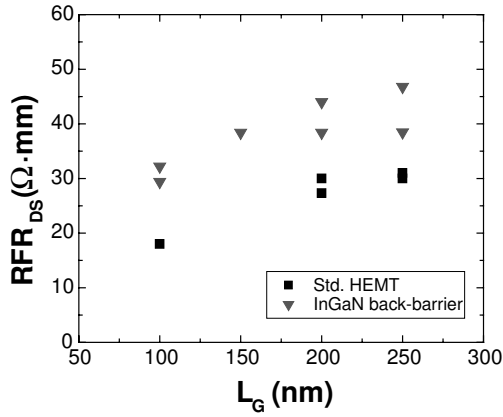


Figure 8.34: RF output resistance in standard AlGaIn/GaN HEMTs and in HEMTs with enhanced confinement due to the use of an InGaIn back-barrier.

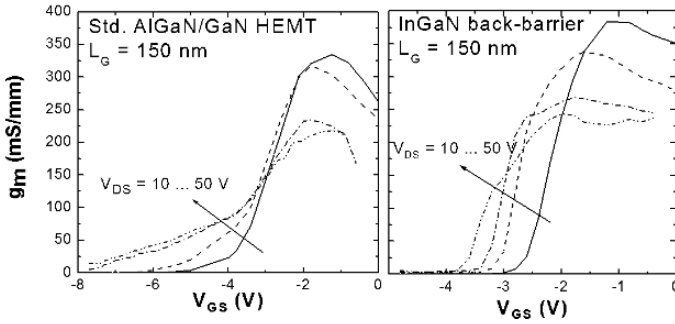


Figure 8.35: Change in  $g_m$  and pinch off with  $V_{DS}$  in a standard AlGaIn/GaN HEMT and in a HEMT with an InGaIn back-barrier.

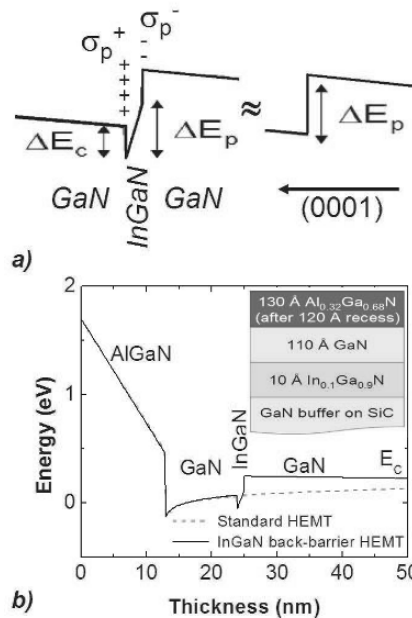


Figure 8.36: a) Effect of the insertion of an ultra-thin layer of InGaN in the conduction band diagram of a GaN buffer. Due to the extremely thin InGaN layer, the conduction band discontinuity,  $\Delta E_c$ , of one side of the heterostructure is canceled by the  $\Delta E_c$  in the other side and it can be neglected, resulting in an effective band discontinuity equal to  $\Delta E_p$ . In the figure, the polarization-induced sheet charges at the heterointerfaces are also shown. b) Schematic and conduction band diagram of the basic InGaN back-barrier sample used in this work.

In conclusion, in this section we have reviewed several of the problems limiting the performance of real HEMTs as well as some of the solutions normally adopted to overcome them. Some solutions, like to keep a good gate aspect ratio, are common to every semiconductor family; others, like the use of InGaN back-barriers, are specific to some materials. Therefore, to fabricate high performance transistors is fundamental not only to understand the physics of the device but also to know the particularities of each material system, its limitations and advanced properties.

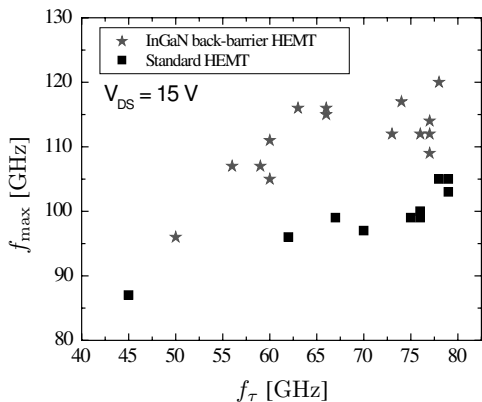


Figure 8.37: Effect of the InGaN back-barrier on the power gain of AlGaIn/GaN HEMTs. Each data point represents a different transistor. The variation in  $f_{\tau}$  is due to different gate lengths, which vary from 0.1 to 0.4  $\mu\text{m}$  for the measured devices.

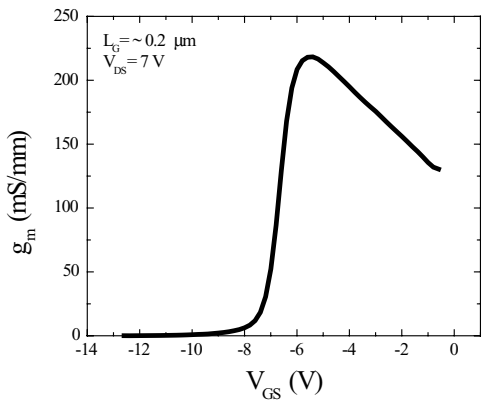


Figure 8.38: Decrease of  $g_m$  as the drain current increases (i.e.  $V_{GS}$  increases) in an AlGaIn/GaN HEMT.

## 8.8 SMALL AND LARGE SIGNAL ISSUES AND FIGURES OF MERIT

It is important to understand the behavior of devices at higher frequencies both in small and large signal operation. The former refers to applications such as low noise amplifiers in receivers whereas the latter to applications such as power amplifiers used in transmitters. In this context several figures of merit have been defined to characterize device performance. It is important to recognize that frequencies of merit are in general a function of the application or equivalently a function of the input and output networks that the device is connected to. In the following sections we will study this in more detail and present in this introduction a short synopsis of the treatment. The most important figure of merit is the current gain cut-off frequency,  $f_T$ , which is proportional to the inverse of the electron transit time across the device. The output termination of the device when  $f_T$  is calculated is always an AC short circuit and hence reflects the device behavior independent of the circuit. The  $f_T$  is the primary indicator of the average electron velocity through the transistor and detailed analysis can extract electron velocity in regions of the transistor. The power gain cutoff frequency of the device  $f_{max}$  is evaluated with the output of the device presented with the complex conjugate of its output impedance to maximize power transfer. This again is predominantly dependent on the device as the termination is determined uniquely by the device characteristic. In other instances, like in large signal amplifiers driving 50 ohms, the load line is what determines the termination and hence another figure of merit,  $f_{lsg}$ , the large signal figure of merit is used. In the discussion of the bipolar device high frequency response we had to discuss minority carrier injection and removal. The FET is a majority carrier device. The device performance is essentially controlled by carrier transit time effects. Thus lithographic techniques defining the gate length and carrier mobility and velocity figure strongly in device response.

### 8.8.1 Small-Signal Characteristics

The equivalent circuit of a MESFET and the source of the various terms are shown in figure 8.39. A change of charge  $\Delta Q$  on the gate produces the change  $\Delta Q$  in the channel (assuming charge neutrality). If  $\Delta t$  is the time taken by the device to respond to this change, the change in the current in the channel is

$$\delta I_D = \frac{\delta Q}{\Delta t} \quad (8.8.1)$$

where  $I_D$  is the current flowing between the source and the drain. The time  $\Delta t$  can be interpreted as the average transit time  $t_{tr}$  for the electrons to move through the device. The transistor transconductance can be related to the transit time. The transistor intrinsic transconductance is

## Appendix D

# IMPORTANT PROPERTIES OF SEMICONDUCTORS

The data and plots shown in this Appendix are extracted from a number of sources. A list of useful sources is given below.

- S. Adachi, J. Appl. Phys., 58, R1 (1985).
- H.C. Casey, Jr. and M.B. Panish, Heterostructure Lasers, Part A, “Fundamental Principles;” Part B, “Materials and Operating Characteristics,” Academic Press, N.Y. (1978).
- Landolt-Bornstein, Numerical Data and Functional Relationship in Science and Technology, Vol. 22, Eds. O. Madelung, M. Schulz, and H. Weiss, Springer-Verlog, N.Y. (1987). Other volumes in this series are also very useful.
- S.M. Sze, Physics of Semiconductor Devices, Wiley, N.Y. (1981). This is an excellent source of a variety of useful information on semiconductors.
- “World Wide Web;” A huge collection of data can be found on the Web. Several professors and industrial scientists have placed very useful information on their websites.

| Material | Structure   | Lattice Constant<br>(Å) | Density<br>(gm/cm <sup>3</sup> ) |
|----------|-------------|-------------------------|----------------------------------|
| C        | Diamond     | 3.5668                  | 3.5153                           |
| Si       | Diamond     | 5.431                   | 2.329                            |
| Ge       | Diamond     | 5.658                   | 5.323                            |
| GaAs     | Zinc Blende | 5.653                   | 5.318                            |
| AlAs     | Zinc Blende | 5.660                   | 3.760                            |
| InAs     | Zinc Blende | 6.058                   | 5.667                            |
| GaN      | Wurtzite    | $a = 3.175; c = 5.158$  | 6.095                            |
| AlN      | Wurtzite    | $a = 3.111; c = 4.981$  | 3.255                            |
| SiC      | Zinc Blende | 4.360                   | 3.166                            |
| Cd       | hcp         | $a = 2.98; c = 5.620$   | 8.65                             |
| Cr       | bcc         | 2.88                    | 7.19                             |
| Co       | hcp         | $a = 2.51; c = 4.07$    | 8.9                              |
| Au       | fcc         | 4.08                    | 19.3                             |
| Fe       | bcc         | 2.87                    | 7.86                             |
| Ag       | fcc         | 4.09                    | 10.5                             |
| Al       | fcc         | 4.05                    | 2.7                              |
| Cu       | fcc         | 3.61                    | 8.96                             |

Table D.1: Lattice constants and density of some semiconductors.

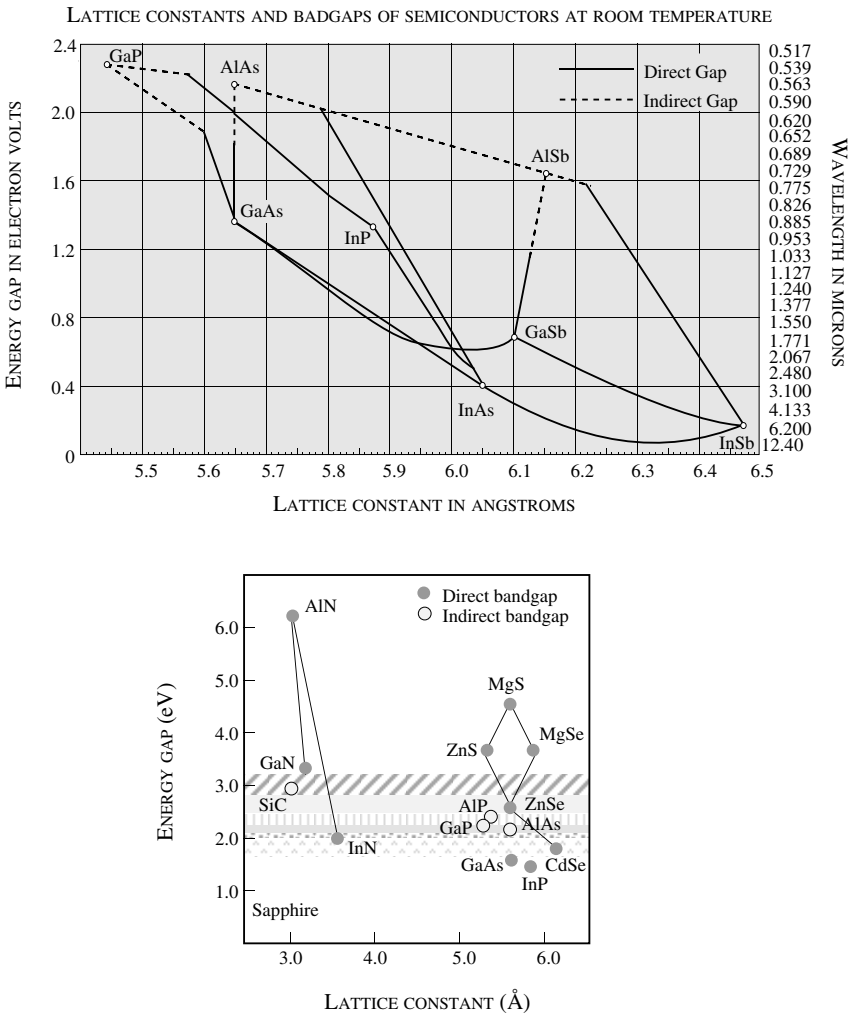


Figure D.1: Lattice constants and bandgaps of semiconductors at room temperature.

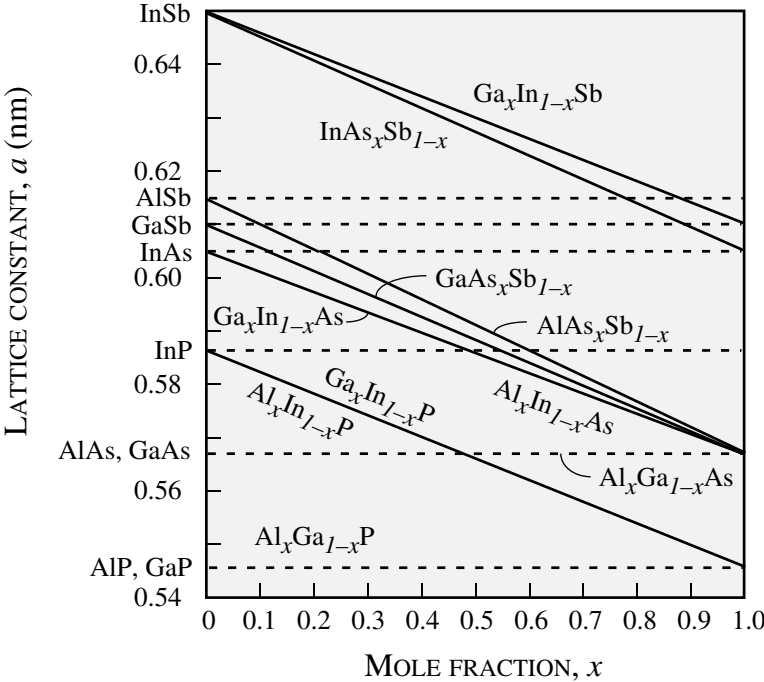


Figure D.2: Lattice constants of several alloy systems.

| Semi-conductor | Type of Energy Gap | Experimental Energy Gap $E_g$ (eV) |       | Temperature Dependence of Energy Gap (eV)      |
|----------------|--------------------|------------------------------------|-------|------------------------------------------------|
|                |                    | 0 K                                | 300 K |                                                |
| AlAs           | Indirect           | 2.239                              | 2.163 | $2.239 - 6.0 \times 10^{-4} T^2 / (T + 408)$   |
| GaP            | Indirect           | 2.338                              | 2.261 | $2.338 - 5.771 \times 10^{-4} T^2 / (T + 372)$ |
| GaAs           | Direct             | 1.519                              | 1.424 | $1.519 - 5.405 \times 10^{-4} T^2 / (T + 204)$ |
| GaSb           | Direct             | 0.810                              | 0.726 | $0.810 - 3.78 \times 10^{-4} T^2 / (T + 94)$   |
| InP            | Direct             | 1.421                              | 1.351 | $1.421 - 3.63 \times 10^{-4} T^2 / (T + 162)$  |
| InAs           | Direct             | 0.420                              | 0.360 | $0.420 - 2.50 \times 10^{-4} T^2 / (T + 75)$   |
| InSb           | Direct             | 0.236                              | 0.172 | $0.236 - 2.99 \times 10^{-4} T^2 / (T + 140)$  |
| Si             | Indirect           | 1.17                               | 1.11  | $1.17 - 4.37 \times 10^{-4} T^2 / (T + 636)$   |
| Ge             | Indirect           | 0.66                               | 0.74  | $0.74 - 4.77 \times 10^{-4} T^2 / (T + 235)$   |

Table D.2: Energy gaps of some semiconductors along with their temperature dependence.

| Material | Electron Mass<br>( $m_0$ )    | Hole Mass<br>( $m_0$ )                            |
|----------|-------------------------------|---------------------------------------------------|
| AlAs     | 0.1                           |                                                   |
| AlSb     | 0.12                          | $m_{\text{dos}} = 0.98$                           |
| GaN      | 0.19                          | $m_{\text{dos}} = 0.60$                           |
| GaP      | 0.82                          | $m_{\text{dos}} = 0.60$                           |
| GaAs     | 0.067                         | $m_{\text{lh}} = 0.082$<br>$m_{\text{hh}} = 0.45$ |
| GaSb     | 0.042                         | $m_{\text{dos}} = 0.40$                           |
| Ge       | $m_l = 1.64$<br>$m_t = 0.082$ | $m_{\text{lh}} = 0.044$<br>$m_{\text{hh}} = 0.28$ |
| InP      | 0.073                         | $m_{\text{dos}} = 0.64$                           |
| InAs     | 0.027                         | $m_{\text{dos}} = 0.4$                            |
| InSb     | 0.13                          | $m_{\text{dos}} = 0.4$                            |
| Si       | $m_l = 0.98$<br>$m_t = 0.19$  | $m_{\text{lh}} = 0.16$<br>$m_{\text{hh}} = 0.49$  |

Table D.3: Electron and hole masses for several semiconductors. Some uncertainty remains in the value of hole masses for many semiconductors.

| Compound                              | Direct Energy Gap<br>$E_g$ (eV) |
|---------------------------------------|---------------------------------|
| $\text{Al}_x\text{In}_{1-x}\text{P}$  | $1.351 + 2.23x$                 |
| $\text{Al}_x\text{Ga}_{1-x}\text{As}$ | $1.424 + 1.247x$                |
| $\text{Al}_x\text{In}_{1-x}\text{As}$ | $0.360 + 2.012x + 0.698x^2$     |
| $\text{Al}_x\text{Ga}_{1-x}\text{Sb}$ | $0.726 + 1.129x + 0.368x^2$     |
| $\text{Al}_x\text{In}_{1-x}\text{Sb}$ | $0.172 + 1.621x + 0.43x^2$      |
| $\text{Ga}_x\text{In}_{1-x}\text{P}$  | $1.351 + 0.643x + 0.786x^2$     |
| $\text{Ga}_x\text{In}_{1-x}\text{As}$ | $0.36 + 1.064x$                 |
| $\text{Ga}_x\text{In}_{1-x}\text{Sb}$ | $0.172 + 0.139x + 0.415x^2$     |
| $\text{GaP}_x\text{As}_{1-x}$         | $1.424 + 1.150x + 0.176x^2$     |
| $\text{GaAs}_x\text{Sb}_{1-x}$        | $0.726 + 0.502x + 1.2x^2$       |
| $\text{InP}_x\text{As}_{1-x}$         | $0.360 + 0.891x + 0.101x^2$     |
| $\text{InAs}_x\text{Sb}_{1-x}$        | $0.18 + 0.41x + 0.58x^2$        |

Table D.4: Compositional dependence of the energy gaps of the binary III-V ternary alloys at 300 K. (After Casey and Panish (1978).)

| Semiconductor                            | Bandgap<br>(eV) |      | Mobility at 300 K<br>(cm <sup>2</sup> /V-s) |       |
|------------------------------------------|-----------------|------|---------------------------------------------|-------|
|                                          | 300 K           | 0 K  | Elec.                                       | Holes |
| C                                        | 5.47            | 5.48 | 1800                                        | 1200  |
| GaN                                      | 3.4             | 3.5  | 1400                                        | 350   |
| Ge                                       | 0.66            | 0.74 | 3900                                        | 1900  |
| Si                                       | 1.12            | 1.17 | 1500                                        | 450   |
| $\alpha$ -SiC                            | 3.00            | 3.30 | 400                                         | 50    |
| GaSb                                     | 0.72            | 0.81 | 5000                                        | 850   |
| GaAs                                     | 1.42            | 1.52 | 8500                                        | 400   |
| GaP                                      | 2.26            | 2.34 | 110                                         | 75    |
| InSb                                     | 0.17            | 0.23 | 80000                                       | 1250  |
| InAs                                     | 0.36            | 0.42 | 33000                                       | 460   |
| InP                                      | 1.35            | 1.42 | 4600                                        | 150   |
| CdTe                                     | 1.48            | 1.61 | 1050                                        | 100   |
| PbTe                                     | 0.31            | 0.19 | 6000                                        | 4000  |
| In <sub>0.53</sub> Ga <sub>0.47</sub> As | 0.8             | 0.88 | 11000                                       | 400   |

Table D.5: Bandgaps along with electron and hole mobilities in several semiconductors. Properties of large bandgap materials (C, GaN, SiC) are continuously changing (mobility is improving), due to progress in crystal growth. Zero temperature bandgap is extrapolated.

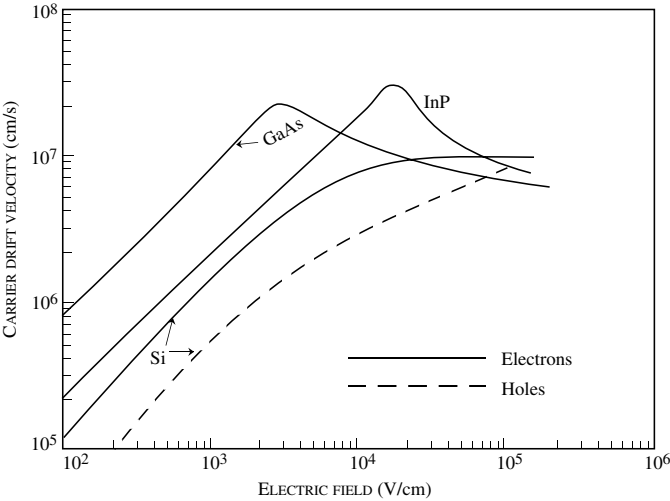


Figure D.3: Velocity-Field relations for several semiconductors at 300 K.

| Material                                 | Bandgap (eV) | Breakdown electric field (V/cm) |
|------------------------------------------|--------------|---------------------------------|
| GaAs                                     | 1.43         | $4 \times 10^5$                 |
| Ge                                       | 0.664        | $10^5$                          |
| InP                                      | 1.34         |                                 |
| Si                                       | 1.1          | $3 \times 10^5$                 |
| In <sub>0.53</sub> Ga <sub>0.47</sub> As | 0.8          | $2 \times 10^5$                 |
| C                                        | 5.5          | $10^7$                          |
| SiC                                      | 2.9          | $2\text{-}3 \times 10^6$        |
| SiO <sub>2</sub>                         | 9            | $\sim 10^7$                     |
| Si <sub>3</sub> N <sub>4</sub>           | 5            | $\sim 10^7$                     |
| GaN                                      | 3.4          | $2 \times 10^6$                 |

Table D.6: Breakdown electric fields in some semiconductors.