Second Edition

Heteroepitaxy of Semiconductors

Theory, Growth, and Characterization



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Second Edition

Heteroepitaxy of Semiconductors

Theory, Growth, and Characterization

John E. Ayers
University of Connecticut, Storrs, USA

Tedi Kujofsa
University of Connecticut, Storrs, USA

Paul Rago
University of Connecticut, Storrs, USA

Johanna E. Raphael
University of Connecticut, Storrs, USA

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2017 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20160321

International Standard Book Number-13: 978-1-4822-5435-8 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Names: Ayers, John E., author. | Kujofsa, Tedi, author. | Rango, Paul, author. | Raphael, Johanna E., author.
Title: Heteroepitaxy of semiconductors : theory, growth, and characterization / John E. Ayers, Tedi Kujofsa, Paul Rango, and Johanna E. Raphael.
Description: Second edition. | Boca Raton : Taylor & Francis Group, a CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa, plc, [2017] | Includes bibliographical references and index.
Identifiers: LCCN 2016013249 | ISBN 9781482254358 (alk. paper)
Subjects: LCSH: Compound semiconductors. | Epitaxy.
Classification: LCC QC611.8.C64 A94 2017 | DDC 537.6/226--dc23
LC record available at http://lccn.loc.gov/2016013249

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface to the Second Edition xiii
Preface to the First Edition xv

1. Introduction 1
References 4

2. Properties of Semiconductors 7
2.1 Introduction 7
2.2 Crystallographic Properties 7
2.2.1 Diamond Structure 12
2.2.2 Zinc Blende Structure 12
2.2.3 Wurtzite Structure 13
2.2.4 Silicon Carbide 14
2.2.5 Miller Indices in Cubic Crystals 15
2.2.6 Miller–Bravais Indices in Hexagonal Crystals 16
2.2.7 Computations and the Metric Tensor 17
2.2.7.1 Coordinate Transformation 18
2.2.7.2 The Metric Tensor 20
2.2.7.3 Distances between Lattice Points and Lengths of Vectors 21
2.2.7.4 Angle between Vector Directions 23
2.2.7.5 Volume of a Unit Cell 24
2.2.7.6 Reciprocal Basis Vectors and Reciprocal Metric Tensor 24
2.2.7.7 Distances and Angles Involving Planes 26
2.2.8 Orientation Effects 29
2.2.8.1 Diamond Semiconductors 29
2.2.8.2 Zinc Blende Semiconductors 29
2.2.8.3 Wurtzite Semiconductors 30
2.2.8.4 Hexagonal Silicon Carbide 32
2.3 Lattice Constants and Thermal Expansion Coefficients 32
2.4 Elastic Properties 37
2.4.1 Infinitesimal Strain Theory 37
2.4.2 Hooke's Law 41
2.4.2.1 Hooke's Law for Isotropic Materials 43
2.4.2.2 Hooke's Law for Cubic Crystals 43
2.4.2.3 Hooke's Law for Hexagonal Crystals 46
2.4.3 Elastic Moduli 46
2.4.3.1 Elastic Moduli for Cubic Crystals 48
2.4.3.2 Elastic Moduli for Hexagonal Crystals 49
2.4.4 Biaxial Stresses and Tetragonal Distortion in Cubic Crystals 50
2.4.5 Biaxial Stresses in Hexagonal Crystals 52
2.4.6 Strain Energy in Cubic Crystals 52
2.4.7 Strain Energy in Nitride Semiconductors 53
2.5 Surface Free Energy 53

2.6 Dislocations 57
2.6.1 Screw Dislocations 58
2.6.2 Edge Dislocations 58
2.6.3 Slip Systems 59
2.6.4 Dislocations in Diamond and Zinc Blende Crystals 61
2.6.4.1 Threading Dislocations in Diamond and Zinc Blende Crystals 63
2.6.4.2 Misfit Dislocations in Diamond and Zinc Blende Crystals 64
2.6.5 Dislocations in Wurtzite Crystals 67
2.6.5.1 Threading Dislocations in Wurtzite Crystals 67
2.6.5.2 Misfit Dislocations in III-Nitrides 69
2.6.6 Dislocations in Hexagonal SiC 69
2.6.6.1 Threading Dislocations in Hexagonal SiC 70
2.6.7 Strain Fields and Line Energies of Dislocations 70
2.6.7.1 Energy of the Screw Dislocation 70
2.6.7.2 Energy of the Edge Dislocation 72
2.6.7.3 Energy of Mixed Dislocations 73
2.6.7.4 Frank's Rule for Dislocation Energies 74
2.6.7.5 Line Energies of Dislocations in Nitride Semiconductors 74
2.6.7.6 Hollow-Core Dislocations (Micropipes) 75
2.6.8 Forces on Dislocations 76
2.6.9 Dislocation Motion 77
2.6.10 Electronic Properties of Dislocations 80
2.6.10.1 Diamond and Zinc Blende Semiconductors 80
2.7 Planar Defects 83
2.7.1 Stacking Faults 83
2.7.2 Twins 85
2.7.3 Inversion Domain Boundaries 86
2.8 Electronic Properties of Semiconductors 88
References 89

3. Heteroepitaxial Growth 101
3.1 Introduction 101
3.2 Vapor Phase Epitaxy 102
3.2.1 Vapor Phase Epitaxial Mechanisms and Growth Rates 102
3.2.2 Hydrodynamic Considerations 104
3.2.3 Vapor Phase Epitaxial Reactors 106
3.2.4 Metalorganic Vapor Phase Epitaxy 109
3.3 Molecular Beam Epitaxy 112
3.4 Silicon, Germanium, and $Si_{1-x}Ge_x$ Alloys 115
3.5 Silicon Carbide 117
3.6 III-Arsenides, III-Phosphides, and III-Antimonides 118
3.7 III-Nitrides 119
3.7.1 Vapor Phase Epitaxial Growth of III-Nitrides 120
3.7.2 Molecular Beam Epitaxy of III-Nitrides 122
3.8 II-VI Semiconductors 123
3.8.1 ZnSe and Its Alloys 124
3.8.2 HgCdTe 124
3.8.3 ZnO 125

3.9 Conclusion 126
References 126

4. Surface and Chemical Considerations in Heteroepitaxy 131
4.1 Introduction 131
4.2 Surface Reconstructions 132
4.2.1 Wood's Notation for Reconstructed Surfaces 134
4.2.2 Experimental Observations 135
4.2.2.1 Si(001) Surface 135
4.2.2.2 Si(111) Surface 136
4.2.2.3 Ge(111) Surface 136
4.2.2.4 6H-SiC(0001) Surface 136
4.2.2.5 3C-SiC(001) 137
4.2.2.6 3C-SiC(111) 137
4.2.2.7 GaN(0001) 138
4.2.2.8 Zinc Blende GaN(001) 138
4.2.2.9 GaAs(001) 138
4.2.2.10 InP(001) 138
4.2.2.11 Sapphire(0001) 139
4.2.3 Surface Reconstruction and Heteroepitaxy 139
4.2.3.1 Inversion Domain Boundaries 139
4.2.3.2 Heteroepitaxy of Polar Semiconductors with Different Ionicities 141
4.3 Nucleation 142
4.3.1 Homogeneous Nucleation 142
4.3.2 Heterogeneous Nucleation 144
4.3.2.1 Macroscopic Model for Heterogeneous Nucleation 145
4.3.2.2 Atomistic Model 146
4.3.2.3 Vicinal Substrates and Step-Flow Growth 149
4.4 Growth Modes 149
4.4.1 Growth Modes in Equilibrium 151
4.4.1.1 Regime I ($f < \varepsilon_1$) 154
4.4.1.2 Regime II ($\varepsilon_1 < f < \varepsilon_2$) 154
4.4.1.3 Regime III ($\varepsilon_2 < f < \varepsilon_3$) 154
4.4.1.4 Regime IV ($f > \varepsilon_3$) 155
4.4.1.5 Equilibrium Growth Modes in SiGe/Si 155
4.4.2 Growth Modes and Kinetic Considerations 155
4.5 Low-Temperature Nucleation Layers 161
4.5.1 Three-Dimensional Growth 161
4.5.2 Surface Roughening 162
4.5.3 Lattice Relaxation 163
4.5.4 Threading Dislocations 163
4.6 Surfactants in Heteroepitaxy 163
4.6.1 Surfactants and Growth Mode 163
4.6.2 Surfactants and Island Shape 165
4.6.3 Surfactants and Misfit Dislocations 165
4.6.4 Surfactants and Ordering in InGaP 166
4.7 Quantum Dots and Self-Assembly 166
4.7.1 Topographically Guided Assembly of Quantum Dots 166

4.7.2 Stressor-Guided Assembly of Quantum Dots ... 169
4.7.3 Vertical Organization of Quantum Dots ... 170
4.7.4 Precision Lateral Placement of Quantum Dots ... 170
References ... 172

5. Mismatched Heteroepitaxial Growth and Strain Relaxation: I. Uniform Layers ... 181
5.1 Introduction ... 181
5.2 Pseudomorphic Growth and the Critical Layer Thickness ... 183
5.2.1 Matthews and Blakeslee Force Balance Model ... 184
5.2.2 Matthews Energy Calculation ... 186
5.2.3 van der Merwe Model ... 190
5.2.4 People and Bean Model ... 191
5.2.5 Effect of the Sign of Mismatch ... 192
5.2.6 Critical Layer Thickness in Islands ... 194
5.2.7 Critical Layer Thickness in Nitride Semiconductors ... 196
5.3 Dislocation Sources ... 201
5.3.1 Homogeneous Nucleation of Dislocations ... 201
5.3.2 Heterogeneous Nucleation of Dislocations ... 203
5.3.3 Dislocation Multiplication ... 204
5.3.3.1 Frank–Read Source ... 205
5.3.3.2 Spiral Source ... 208
5.3.3.3 Hagen–Strunk Multiplication ... 211
5.4 Interactions between Misfit Dislocations ... 212
5.5 Lattice Relaxation Mechanisms ... 214
5.5.1 Bending of Substrate Dislocations ... 214
5.5.2 Glide of Half Loops ... 216
5.5.3 Injection of Edge Dislocations at Island Boundaries ... 218
5.5.4 Nucleation of Shockley Partial Dislocations ... 219
5.5.5 Cracking ... 221
5.5.6 Interfacial Misfit Dislocation Growth Mode ... 221
5.6 Quantitative Models for Lattice Relaxation ... 225
5.6.1 Matthews and Blakeslee Equilibrium Model ... 226
5.6.2 Kinetic Models for Relaxation ... 227
5.6.3 Dislocation Blocking ... 229
5.6.4 Surface Roughness and Dislocation Blocking ... 232
5.6.5 Matthews, Mader, and Light Kinetic Model ... 234
5.6.6 Dodson and Tsao Kinetic Model ... 235
5.6.7 Hull, Bean, and Buescher Kinetic Model ... 237
5.6.8 Kujofsa et al. Kinetic Model ... 241
5.6.9 Kinetically Limited Lattice Relaxation in Zinc Blende Semiconductors ... 242
5.6.9.1 Lattice Mismatch and Thickness Dependence of Kinetically Limited In-Plane Strain ... 243
5.6.9.2 Temperature-Graded Heterostructures ... 243
5.6.9.3 Lattice Relaxation in the InGaAs/GaAs Material System ... 244
5.6.10 Kinetically Limited Relaxation in Nitride Semiconductors ... 252
5.7 Lattice Relaxation on Vicinal Substrates: Crystallographic Tilting of Heteroepitaxial Layers ... 254
5.7.1 Nagai Model ... 254

5.7.2 Olsen and Smith Model 255
5.7.3 Ayers, Ghandhi, and Schowalter Model 256
5.7.4 Riesz Model 263
5.7.5 Vicinal Epitaxy of III-Nitride Semiconductors 266
5.7.6 Vicinal Heteroepitaxy with a Change in Stacking Sequence 268
5.7.7 Vicinal Heteroepitaxy with Multilayer Steps 269
5.8 Dislocation Coalescence, Annihilation, and Removal in Relaxed Heteroepitaxial Layers 271
5.8.1 Thermal Strain 275
5.8.2 Cracking in Thick Films 277
References 280

6. Mismatched Heteroepitaxial Growth and Strain Relaxation: II. Graded Layers and Multilayered Structures 289
6.1 Introduction 289
6.2 Critical Layer Thickness: General Case 289
6.3 Equilibrium Strain and Misfit Dislocations: General Case 292
6.4 Kinetically Limited Strain Relaxation: General Case 294
6.5 Threading Dislocation Densities: General Case 297
6.6 Step-Graded Layer 299
6.6.1 Lattice Relaxation and Residual Strain in a Step-Graded Layer 300
6.6.2 Misfit and Threading Dislocations in a Step-Graded Layer 303
6.6.3 Morphology and Surface Roughening in a Step-Graded Layer 306
6.6.4 Crystallographic Tilting in a Step-Graded Layer 308
6.7 Linearly Graded Layer 308
6.7.1 Approaches to Linear Grading 308
6.7.2 Critical Thickness in a Linearly Graded Layer 309
6.7.3 Critical Layer Thicknesses in Linearly Graded Layers with Nonzero Interfacial Mismatch 311
6.7.4 Misfit Dislocations and Strain in a Linearly Graded Layer 317
6.7.5 Threading Dislocations in a Linearly Graded Layer 323
6.7.6 Crystallographic Tilting in a Linearly Graded Layer 330
6.7.7 Surface Roughening and Cross-Hatch in a Linearly Graded Layer 333
6.7.8 Dual-Slope and Tandem Graded Layers 333
6.8 Sublinearly and Superlinearly Graded Layers 335
6.8.1 Critical Layer Thickness in Sublinear Exponentially Graded Layers 336
6.8.2 Strain Relaxation in Sublinearly Graded Layers 342
6.8.3 Comparison of Sublinearly and Superlinearly Graded Layers 342
6.9 S-Graded Layer 346
6.9.1 Misfit Dislocations and Strain in the S-Graded Layer 346
6.9.2 Refined Model for S-Graded Layers 354
6.9.3 Threading Dislocations in S-Graded Layers 355
6.10 Strained Layer Superlattices 357
6.11 Conclusion 358
References 359

7. Characterization of Heteroepitaxial Layers 367
7.1 Introduction 367
7.2 X-Ray Diffraction 368

7.2.1 Positions of Diffracted Beams 369
7.2.1.1 Bragg Equation 369
7.2.1.2 Reciprocal Lattice and the von Laue Formulation for Diffraction 370
7.2.1.3 Ewald Sphere 372
7.2.2 Intensities of Diffracted Beams 372
7.2.2.1 Scattering of X-Rays by a Single Electron 373
7.2.2.2 Scattering of X-Rays by an Atom 374
7.2.2.3 Scattering of X-Rays by a Unit Cell 375
7.2.2.4 Intensities of Diffraction Profiles 377
7.2.3 Dynamical Diffraction Theory 377
7.2.3.1 Intrinsic Diffraction Profiles for Perfect Crystals 378
7.2.3.2 Intrinsic Widths of Diffraction Profiles 380
7.2.3.3 Extinction Depth and Absorption Depth 381
7.2.4 X-Ray Diffractometers 382
7.2.4.1 Double-Crystal Diffractometer 383
7.2.4.2 Bartels Double-Axis Diffractometer 386
7.2.4.3 Triple-Axis Diffractometer 387
7.2.5 Resolution of X-Ray Diffraction Measurements and the Effect of Finite Counting Statistics 387
7.2.6 Reciprocal Space Maps 389
7.3 Electron Diffraction 392
7.3.1 Reflection High-Energy Electron Diffraction 392
7.3.2 Low-Energy Electron Diffraction 394
7.4 Microscopy 395
7.4.1 Optical Microscopy 395
7.4.2 Transmission Electron Microscopy 396
7.4.3 Scanning Tunneling Microscopy 398
7.4.4 Atomic Force Microscopy 400
7.5 Crystallographic Etching Techniques 400
7.6 Photoluminescence 402
7.7 Growth Rate and Layer Thickness 405
7.8 Determination of Composition and Strain 408
7.8.1 X-Ray Diffraction Analysis of a Binary Heteroepitaxial Layer 408
7.8.2 X-Ray Diffraction Analysis of a Ternary Heteroepitaxial Layer 410
7.8.3 X-Ray Analysis of a Quaternary Heteroepitaxial Layer 414
7.8.4 Analysis of Composition and Strain Using X-Ray Reciprocal Space Maps 414
7.8.4.1 Application of RSM to a Uniform Buffer Layer 414
7.8.4.2 Application of RSM to Linearly Graded and Step-Graded Buffers 417
7.8.4.3 Application of Reciprocal Space Maps to III-Nitride Materials 418
7.8.5 *In Situ* Stress–Strain Measurements Using Multibeam Optical Stress Sensor 419
7.9 Determination of the Critical Layer Thickness 424
7.9.1 Effect of Finite Resolution on CLT Determination 425
7.9.2 X-Ray Diffraction Determination of the CLT 427

7.9.2.1 X-Ray Strain Method for CLT Determination 428
7.9.2.2 X-Ray FWHM Method for CLT Determination 432
7.9.2.3 X-Ray Topography Determination of CLT 437
7.9.2.4 TEM Determination of the CLT 438
7.9.2.5 Electron Beam–Induced Current Determination of CLT 439
7.9.2.6 Photoluminescence Determination of the CLT 439
7.9.2.7 PLM Determination of the CLT 441
7.9.2.8 RHEED Determination of the CLT 442
7.9.2.9 Scanning Tunneling Microscope Determination of the CLT 444
7.9.2.10 Rutherford Backscattering Determination of the CLT 445
7.10 Determination of the Crystal Orientation 447
7.11 Characterization of Defect Types and Densities 449
7.11.1 Characterizing Defects by TEM 450
7.11.2 Characterizing Defects by Crystallographic Etching 451
7.11.3 Characterizing Defects by X-Ray Diffraction 453
7.11.4 Characterization of Asymmetric Dislocation Densities 459
7.11.5 Diffuse Scattering Characterization of Defects 463
7.11.5.1 Application of X-Ray Reciprocal Space Maps to Characterize Defects in IMF-Grown Materials 463
7.12 Dynamical X-Ray Diffraction Analysis of Multilayered Device Structures and Superlattices 465
7.12.1 Dynamical X-Ray Diffraction Analysis of Device Heterostructures Containing Dislocations 469
7.12.1.1 Phase-Invariant Model 470
7.12.1.2 Dynamical Modeling of Asymmetrical Dislocation Densities 480
7.12.1.3 Mosaic Crystal Model 484
7.13 Characterization of the Growth Mode 490
References 494

8. Defect Engineering in Heteroepitaxial Material 503
8.1 Introduction 503
8.2 Buffer Layer Approaches and Virtual Substrates 503
8.2.1 Dislocation Compensation 504
8.2.2 Temperature-Graded Layers 507
8.2.3 Superlattice Buffer Layers 507
8.3 Reduced Area Growth Using Patterned Substrates 513
8.4 Patterning and Annealing 516
8.5 Defect Reduction by Selective Evaporation 521
8.6 Epitaxial Lateral Overgrowth 523
8.7 Pendeo-Epitaxy 529
8.8 Nanoheteroepitaxy 529
8.8.1 Nanoheteroepitaxy on a Noncompliant Substrate 532
8.8.2 Nanoheteroepitaxy with a Compliant Substrate 534
8.9 Planar Compliant Substrates 537
8.9.1 Compliant Substrate Theory 539
8.9.2 Compliant Substrate Implementation 541

8.9.2.1 Cantilevered Membranes 543
8.9.2.2 Silicon-on-Insulator as a Compliant Substrate 544
8.9.2.3 Twist-Bonded Compliant Substrates 548
8.10 Free-Standing Semiconductor Films 551
8.11 Conclusion 552
References 553

9. Metamorphic Devices 557
9.1 Introduction 557
9.2 Strain-Relaxed Buffer MOSFETs 558
9.3 High-Electron-Mobility Transistors 560
9.3.1 HEMTs with Arsenide Channel Layers 560
9.3.2 HEMTs with Nitride Channel Layers 561
9.4 Heterojunction Bipolar Transistors 563
9.5 Light-Emitting Diodes 564
9.5.1 Red Light-Emitting Diodes 566
9.5.2 Amber Light-Emitting Diodes 570
9.5.3 Green Light-Emitting Diodes 571
9.5.4 Blue Light-Emitting Diodes 573
9.5.5 Ultraviolet Light-Emitting Diodes 575
9.5.6 White Light-Emitting Diodes 577
9.6 Solar Cells 579
9.7 Conclusions 586
References 586

Appendix I 593
Appendix II 595
Appendix III 597
Appendix IV 603
Appendix V 609
Appendix VI 611
Appendix VII 613
Appendix VIII 617
Index 629

Preface to the Second Edition

Since the printing of the first edition, heteroepitaxy has only increased in importance with the explosive growth of the electronics industry and the development of a myriad of heteroepitaxial devices for solid-state lighting, green energy, displays, communications, and digital computing. Our ever-growing understanding of the basic physics and chemistry underlying heteroepitaxy, and especially lattice relaxation and dislocation dynamics, has enabled an ever-increasing emphasis on metamorphic devices, creating great freedom in the choice of semiconductor materials for the many applications named above. To reflect this emphasis, we have included two all-new chapters that address metamorphic buffer layers (Chapter 6) and metamorphic devices (Chapter 9). The other seven chapters have been revised extensively, with new material on crystal symmetry and relationships, III-nitride materials, lattice relaxation physics and models, *in situ* characterization, and reciprocal space maps. The basic approach remains the same, however, with a focus on principles that should remain useful for some time to come. It is our hope that this approach, and the extensive new material, will make the second edition an even more valuable reference for researchers in the field.

John E. Ayers
Tedi Kujofsa
Paul B. Rago
Johanna Raphael
Storrs, Connecticut



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Preface to the First Edition

Heteroepitaxy, or the single-crystal growth of one semiconductor on another, has been the topic of intense research for several decades. This effort received a significant boost with the advent of MOVPE, MBE, and other advancements in epitaxial growth. It became possible to grow almost any semiconductor material or structure, including alloys, multilayers, superlattices, and graded layers, with unprecedented control and uniformity. Researchers embraced these capabilities and set out to grow nearly every imaginable combination of epitaxial layer/substrate. Across this great diversity of materials and structures, there has begun to emerge a general understanding of at least some aspects of heteroepitaxy, especially nucleation, growth modes, relaxation of strained layers, and dislocation dynamics. The application of this knowledge has now enabled the commercial production of a wide range of heteroepitaxial devices, including high-brightness light-emitting diodes, lasers, and high-frequency transistors, to name a few.

Our understanding of heteroepitaxy is far from complete, and the field is evolving rapidly. Here, I did not attempt to report all of the results from every known heteroepitaxial material combination. Even if this had been possible, such a book would become out of date with the next wave of electronic journals. Instead, I tried to emphasize the principles underlying heteroepitaxial growth and characterization, with many examples from the material systems which have been studied. I hope that this approach will remain useful for some time to come, as a reference to researchers in the field and also as a starting point for graduate students.

I would like to express my sincere gratitude to Professor Sorab K. Ghandhi, who introduced me to the field of heteroepitaxy. I am also indebted to my graduate students and my fellow researchers, without whom this book would not be possible. Finally, I would like to thank my family for their unending support and patience throughout this endeavor.

John E. Ayers
Storrs, Connecticut



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

1

Introduction

Modern semiconductor devices require materials with a diversity of lattice constants, thermal expansion coefficients, and even crystal orientation or structure. Some of these materials are shown in the "bandgap engineering" diagrams of Figures 1.1 and 1.2. Because of the wide range of semiconductor materials used in devices, and the limited choices of single-crystal substrates on which to fabricate them, their manufacture almost always requires lattice-mismatched heteroepitaxial growth.[1] Initial efforts were concentrated on pseudomorphic growth,[2] in which the epitaxial materials are coherently strained and absent of misfit dislocations. However, the requirement to avoid lattice relaxation places considerable restrictions on the choices of compositions and layer thicknesses.[3] In general, the thickness mismatch product is restricted to less than ~0.04 nm for individual layers within pseudomorphic structures.[4] Mismatched epitaxy brings with it a host of challenges, including strain,[5] misfit dislocations,[2] and the associated threading dislocations that propagate through device regions,[6] crystallographic tilt induced during relaxation,[7] and the possible degradation of morphology due to three-dimensional nucleation[8] or stress-induced surface roughening.[9] Foremost among these problems is the density of threading dislocations, which can exceed 10^9 cm^{-2} and strongly influences the performance and reliability of devices utilizing the defected material.[10]

The benefits of metamorphic heteroepitaxy greatly outweigh the challenges outlined above, however. First, heteroepitaxy enables semiconductor devices of nearly any composition, without the need for the identification of suitable lattice-matched substrates. It is important to note that even in lattice-matched devices, such as $(Al_xGa_{1-x})_{0.52}In_{0.48}P$ light-emitting diodes on GaAs substrates,[11] differences in thermal expansion coefficients as well as limitations in compositional control give rise to lattice mismatch and dislocations. Second, heteroepitaxy enables the fabrication of heterojunction devices, such as lasers, light-emitting diodes (LEDs), and high-electron-mobility transistors (HEMTs). Third, devices made from different semiconductors may be integrated onto the same wafer, as in multijunction solar cells[12] or the integration of optoelectronics and very large-scale integrated (VLSI) circuits.

The purpose of this book is to describe the principles involved in the heteroepitaxy of semiconductors and give some important examples of its application to metamorphic devices. The following chapters review the properties of semiconductors, heteroepitaxial growth methods, surface and chemical considerations, mismatched heteroepitaxy and strain relaxation, characterization of heteroepitaxial layers, defect engineering approaches, and metamorphic devices.

The strain relaxation, dislocation dynamics, and defect densities in heteroepitaxial layers are strongly influenced by the lattice constants, elastic constants, thermal expansion behavior, dislocation energetics, and kinetics. Chapter 2 reviews fundamental properties of semiconductors that bear on these aspects of mismatched heteroepitaxy.

Many advancements in the field of heteroepitaxy would not have been possible without the development of the epitaxial growth techniques—molecular beam epitaxy (MBE) and metalorganic vapor phase epitaxy (MOVPE)—which are reviewed in Chapter 3. These

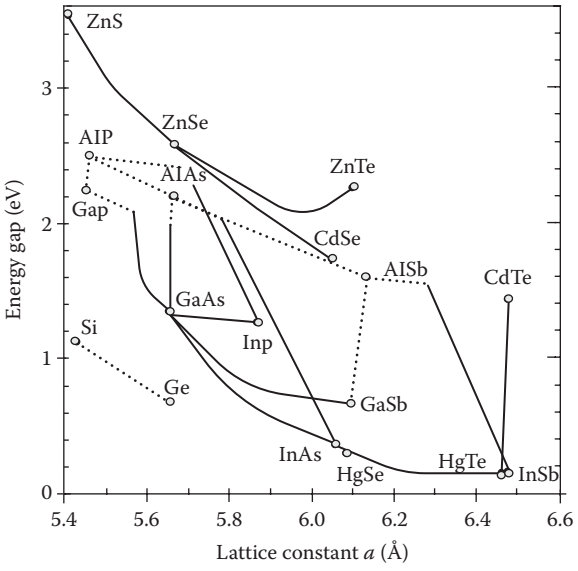


FIGURE 1.1
Energy gap as a function of lattice constant *a* for cubic semiconductors. Room temperature values are given. Dashed lines indicate an indirect gap.

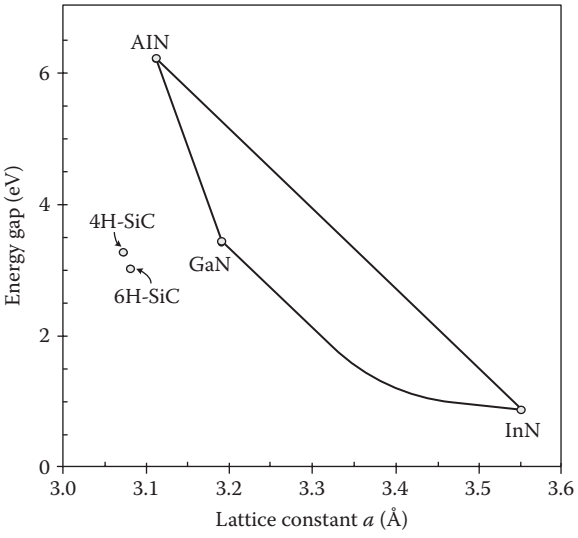


FIGURE 1.2
Energy gap as a function of lattice constant *a* for hexagonal semiconductors. Room temperature values are given. Sapphire, a commonly used substrate material for III-nitrides, has room temperature lattice constants of $a = 4.7592$ Å and $c = 12.9916$ Å. (From Y. V. Shvyd'ko et al., *J. Synchrotron Rad.*, 9, 17, 2002.)

methods afford tremendous flexibility and the ability to deposit thin layers and complex multilayered structures with precise control and excellent uniformity. In addition, the high-vacuum environment of MBE makes it possible to employ *in situ* characterization tools using electron and ion beams, which provide the crystal grower with immediate feedback, and improved control of the growth process. For these reasons, MBE and MOVPE have

emerged as general-purpose tools for heteroepitaxial research and commercial production. Together, these two epitaxial growth methods account for virtually all production of compound semiconductor devices today.

An important distinction between heteroepitaxy and homoepitaxy is the need to nucleate a new phase on the substrate surface. Therefore, the surface and its structure, as well as surface-segregated impurities (surfactants), can play important roles in determining the usefulness of heteroepitaxial layers for the fabrication of devices. Chapter 4 provides an in-depth description of semiconductor crystal surfaces and their reconstructions, nucleation, growth modes, and use of surfactants. Control of the growth mode, through the tailoring of growth conditions or the use of surfactants, is believed to be critical to the development of some devices. Two-dimensional growth is desirable in most cases, for the achievement of flat, abrupt interfaces and surfaces, and is mandated for quantum well devices. For the development of quantum dot devices, Volmer–Weber (island growth) or Stranski–Krastanov (growth of a continuous wetting layer followed by islanding) growth is actually desirable. Here, the control of the sizes, shapes, and distributions of islands is critical. This aspect, called the self-assembly of quantum dots, is also discussed in Chapter 4.

Heteroepitaxial growth is often metamorphic, so strain relaxation and the associated creation of crystal defects are of great importance. Under the condition of low lattice mismatch (<2%) it is possible to grow a pseudomorphic heteroepitaxial layer, which maintains coherency with the substrate crystal in the plane of the interface. At some thickness (the critical layer thickness), the creation of misfit dislocations becomes energetically favorable for relaxing the excess stress associated with the growth of mismatched materials. The lattice relaxation process is rather complex, and is usually limited by the nucleation, multiplication, glide, or climb of dislocations. Often, gliding dislocations are blocked by dislocations with orthogonal misfit components, depending on the strain and film thickness. Invariably, nonequilibrium threading dislocations are introduced, together with the stress-relieving misfit defects, and their concentration is determined by the aforementioned physical processes. The presence of dislocations in the material tends to degrade its electronic properties, affecting device performance and lifetime. The control of the lattice relaxation process is therefore an area of considerable interest. Chapter 5 provides an in-depth review of mismatched heteroepitaxy and lattice relaxation in uniform layers, while Chapter 6 considers graded and multilayered structures.

Characterization tools have played a key role in the advancement of the science of heteroepitaxy. Some of the most commonly used techniques are microscopic techniques, x-ray diffraction, photoluminescence, and crystallographic etching. These are covered in detail in Chapter 7, with an emphasis on x-ray diffraction, which is the most widely used nondestructive tool for structural characterization of heteroepitaxial layers. Individual sections are also devoted to some key application areas for these characterization tools, such as the determination of the stress, strain, and composition; the determination of the critical layer thickness for lattice relaxation; the characterization of the morphology and growth mode; the observation of crystal defects; and the determination of their types and configurations.

The broad application of mismatched heteroepitaxy to device and circuit fabrication requires the control of the crystal defect structures, and therefore a number of defect engineering approaches have emerged. Of these, buffer layer approaches have been broadly applied and are discussed in detail in Chapter 6. Chapter 8 reviews other defect engineering approaches, including reduced area growth, dislocation compensation, patterning and annealing, epitaxial lateral overgrowth (ELO), nanoheteroepitaxy, and compliant

substrates. All of these were designed to reduce the dislocation densities of heteroepitaxial layers to practical levels for device applications. Some are intended to remove existing defects from lattice-relaxed heteroepitaxial layers, such as patterning and annealing, ELO, or superlattice buffer layers (dislocation filters). Others were conceived in order to prevent lattice relaxation in the first place; these include reduced area growth, nanoheteroepitaxy, and compliant substrates.

The proliferation of defect engineering methods could be taken as an indication that none of them are uniquely suited to the purpose, for all material systems. On the other hand, some of these approaches have been highly successful, to the point of being used in commercial devices. Graded buffer layers are the most important example of this, and have been used in commercial $GaAs_{1-x}P_x$ LEDs on GaAs substrates and $In_xGa_{1-x}As$ HEMTs on GaAs substrates. ELO is an important method used to reduce the threading dislocation densities in the active regions of III-nitride lasers. Other defect engineering approaches, such as the use of compliant substrates, show great promise but have not yet found commercial use. In order to tap the great potential of heteroepitaxy, defect engineering approaches will continue to be important, not only in the applications listed above but also in new ones.

Chapter 9 discusses some of the important applications of metamorphic materials to semiconductor devices, including strain-relaxed buffer (SRB) metal oxide semiconductor field effect transistors (MOSFETs), HEMTs, LEDs, and solar cells. This list, though not exhaustive, provides a good cross section of applications and illustrates the relationships between processing, defects, strain, and surface roughness, and device characteristics. The rapid progress in the development of metamorphic devices can be expected to continue and, in fact, accelerate, based on the great success already achieved in this area.

References

1. E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, Dislocation dynamics in relaxed graded composition semiconductors, *Mater. Sci. Eng.*, B67, 53 (1999).
2. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers. I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).
3. W. E. Hoke, T. D. Kennedy, A. Torabi, C. S. Whelan, P. F. Marsh, R. E. Leoni, S. M. Lardizabal, et al., Properties of metamorphic materials and device structures on GaAs substrates, *J. Cryst. Growth*, 251, 804 (2003).
4. L. J. Mawst, J. D. Kirch, C.-C. Chang, T. Kim, T. Garrod, D. Botez, S. Ruder, et al., InGaAs/AlInAs strain-compensated superlattices grown on metamorphic buffer layers for low-strain, 3.6 μm-emitting quantum-cascade-laser active regions, *J. Cryst. Growth*, 370, 230 (2013).
5. J. H. van der Merwe, Crystal interfaces. Part II. Finite overgrowths, *J. Appl. Phys.*, 34, 123 (1963).
6. M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, Dislocation morphology in graded heterojunctions: $GaAs_{1-x}P_x$, *J. Mater. Sci.*, 4, 223 (1969).
7. G. H. Olsen and R. T. Smith, Misorientation and tetragonal distortion in heteroepitaxial vapor-grown III-V structures, *Phys. Status Solidi A*, 31, 739 (1975).
8. I. Daruka and A.-L. Barabasi, Dislocation-free island formation in heteroepitaxial growth: A study at equilibrium, *Phys. Rev. Lett.*, 79, 3708 (1997).
9. J. Tersoff, Stress-induced roughening in epitaxial growth, *Appl. Surf. Sci.*, 102, 1 (1996).
10. A. E. Romanov, W. Pompe, S. Mathis, G. E. Beltz, and J. S. Speck, Threading dislocation reduction in strained layers, *J. Appl. Phys.*, 85, 182 (1999).

11. M. R. Krames, O. R. Shchkekin, R. Mueller-Mach, G. O. Mueller, L. Zhou, G. Harbers, and M. G. Craford, Status and future of high-power light-emitting diodes for solid state lighting, *J. Display Technol.*, 3, 160 (2007).
12. M. Stan, D. Aiken, B. Cho, A. Cornfeld, V. Ley, P. Patel, P. Sharps, and T. Varghese, High-efficiency quadruple junction solar cells using OMVPE with inverted metamorphic device structures, *J. Cryst. Growth*, 312, 1370 (2010).
13. Y. V. Shvyd'ko, M. Lucht, E. Gerdau, M. Lerche, E. E. Alp, W. Sturhahn, J. Sutter, and T. S. Toellner, Measuring wavelengths and lattice constants with the Mössbauer wavelength standard, *J. Synchrotron Rad.*, 9, 17 (2002).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

2

Properties of Semiconductors

2.1 Introduction

The key challenges in the heteroepitaxy of semiconductors, relative to the development of useful devices, include the control of the growth morphology, stress and strain, and crystal defects. The purpose of this chapter is to review the properties of semiconductors that bear on these aspects of heteroepitaxy, including crystallographic properties, elastic properties, surface properties, and defect structures.

2.2 Crystallographic Properties

Semiconductors that are commonly in use today are nearly always single-crystal materials.* A crystal is a periodic arrangement of atoms in space. A space lattice and a basis comprise a crystal structure. The space lattice describes the periodic arrangement of points on which atoms (or groups of atoms) may be placed, whereas the basis can be a single atom or an arrangement of atoms placed at each space lattice point. A basis consisting of a single atom is monoatomic, whereas a basis consisting of multiple atoms is polyatomic.

Considering Haüy's block model, a three-dimensional (3-D) crystal can be constructed by the stacking of a particular unit cell. An example of this is shown in Figure 2.1. The unit cell illustrated in Figure 2.1 can be described by choosing one of the corners as the origin and defining three vectors, $\vec{a}$, $\vec{b}$, and $\vec{c}$, along the adjacent edges. These are known as the basis vectors, and they may be used to identify any of the other lattice points in the structure through the use of a translation vector $\vec{t}$.

The resulting collection of points gives rise to the space lattice, with each corner forming a node or lattice point. The translation vector defines the lattice displacement from the origin by linear combinations of the three basis vectors. In addition, if we were to translate the entire lattice using a translation vector $\vec{t}$, we would obtain an invariant (undistinguishable) lattice from the one prior to translation. This property is known as translational symmetry; the environment is the same when viewed from two distinct points separated by a translation vector $\vec{t}$.

Translation from the origin to a lattice point with coordinates (u, v, w) requires a translation vector $\vec{t}$ given by

$$\vec{t} = u\vec{a} + v\vec{b} + w\vec{c}, \tag{2.1}$$

* Notable exceptions include thin-film transistors, made using polycrystalline or amorphous silicon, and the gates of metal oxide semiconductor field effect transistors, which have traditionally been made using polycrystalline silicon.

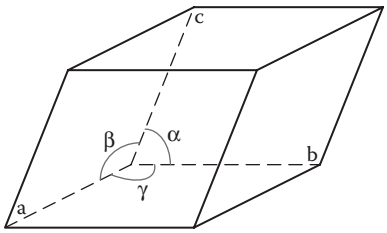


FIGURE 2.1
General 3-D unit cell with lattice parameters.

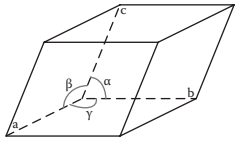
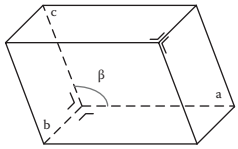
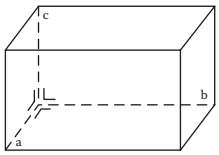
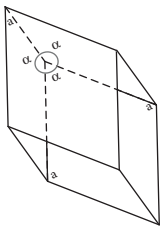
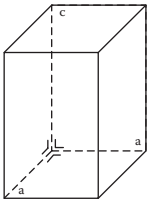
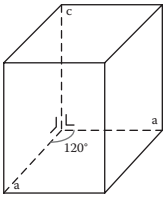
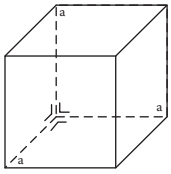
where u, v, and w are integers. In order to shorten the vector notation, we refer to vector quantities such as $\vec{t}$ by a boldface symbol, **t**. In principle, the set of translation vectors of the space lattice contains an infinite number of elements, although real crystals have finite size.

The parameters necessary to describe a unit cell are given by $\{a, b, c, \alpha, \beta, \gamma\}$, where a, b, and c are the magnitude of the basis vectors and the angles α, β, and γ are the angles between the basis vectors $\vec{b}$ and $\vec{c}$, $\vec{a}$ and $\vec{c}$, and $\vec{a}$ and $\vec{b}$, respectively. The lattice parameters are not constant for real crystals and their values depend on temperature, pressure, and composition, but we will treat them as constants in sections to come. The consideration of the six parameters yields seven fundamentally different 3-D crystal structures known as the primitive cells, and a summary of them is given in Table 2.1. The seven primitive 3-D crystals are triclinic (sometimes referred to as anorthic), monoclinic, orthorhombic, rhombohedral (sometimes referred to as trigonal), tetragonal, hexagonal, and cubic. Table 2.1 lists the crystal structures according to their characteristic symmetry where the minimum symmetry is governed by rotations that are compatible with translation. In addition, the metric tensors that will be explained in later sections are given to allow an indirect correlation to the symmetric elements of each primitive crystal structure. In addition, Table 2.1 gives restrictions (if any) on the angles between the basis vectors. A primitive cell is regarded as a minimum volume cell, and it contains one lattice point per unit cell. In general, one describes a crystal lattice using the shortest combination of the basis vectors (primitive cell). However, there also exists another choice of basis vectors that describes the minimum volume cell, known as the Wigner–Seitz (WS) cell, and this reveals the true symmetry of the lattice. The basis vector choice to describe such a cell is not as straightforward as in the previous cases because a WS cell may contain anywhere between 6 and 14 faces. A WS cell is commonly used in solid-state physics to study quantum mechanics of minimum volume cells. Although it may be possible to describe any crystal structure using a primitive cell, the use of simpler basis vectors such as the case of nonprimitive cells enables a better distinction of the *cubic* symmetry characteristics of the crystal structure. Therefore, it may also be possible to add additional points to a space lattice without changing the crystal symmetry. However, it requires that the additional lattice points are centered with respect to the original lattice nodes. The three possible types of centering are body centering (BC), face centering (FC), and base centering. The inclusion of these types of centering translations to the primitive cell gives rise to seven new crystal structures; these crystal structures are known as the nonprimitive cells, and they contain more than one lattice point per unit cell. In the case of body centering, the centering vector is given by

$$\mathbf{I} = \frac{\mathbf{a} + \mathbf{b} + \mathbf{c}}{2}, \tag{2.2}$$

TABLE 2.1

Summary of the 3-D Crystal Systems along with Their Metric Tensor and Symmetric Characteristics

Crystal System {Lattice Parameters} Requirements	**3-D Structural Representation of Primitive Cell**	**Metric Tensor**	**Minimum Symmetric Characterization**
Triclinic $\{a,b,c,\alpha,\beta,g\}$ No Restrictions		$\begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ba\cos\gamma & b^2 & bc\cos\alpha \\ ca\cos\beta & cb\cos\alpha & c^2 \end{bmatrix}$	None
Monoclinic $\{a,b,c,90°,\beta,90°\}$ $\alpha=\gamma=90°$		$\begin{bmatrix} a^2 & 0 & ac\cos\beta \\ 0 & b^2 & 0 \\ ca\cos\beta & 0 & c^2 \end{bmatrix}$	One two-fold Axis
Orthorhombic $\{a,b,c,90°,90°,90°\}$ $\alpha=\beta=\gamma=90°$		$\begin{bmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}$	Three mutually perpendicular two-fold axes
Rhombohedral $\{a,a,a,\alpha,\alpha,\alpha\}$ $a=b=c$ $\alpha=\beta=\gamma$		$\begin{bmatrix} a^2 & a^2\cos\alpha & a^2\cos\alpha \\ a^2\cos\alpha & a^2 & a^2\cos\alpha \\ a^2\cos\alpha & a^2\cos\alpha & a^2 \end{bmatrix}$	One three-fold axis
Tetragonal $\{a,a,c,90°,90°,90°\}$ $a=b$ $\alpha=\beta=\gamma=90°$		$\begin{bmatrix} a^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}$	One four-fold axis
Hexagonal $\{a,a,c,90°,90°,120°\}$ $a=b$ $\alpha=\beta=90°$ $\gamma=120°$		$\begin{bmatrix} a^2 & -a^2/2 & 0 \\ -a^2/2 & a^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}$	One six-fold axis
Cubic $\{a,a,a,90°,90°,90°\}$ $a=b=c$ $\alpha=\beta=\gamma=90°$		$\begin{bmatrix} a^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2 \end{bmatrix}$	Four three-fold axes

whereas in the case of face centering or base centering, a node can be added between two existing lattice points using one of the following lattice vectors,

$$\mathbf{A} = \frac{\mathbf{b} + \mathbf{c}}{2}, \tag{2.3}$$

$$\mathbf{B} = \frac{\mathbf{a} + \mathbf{c}}{2}, \tag{2.4}$$

$$\mathbf{C} = \frac{\mathbf{a} + \mathbf{b}}{2}. \tag{2.5}$$

The combination of primitive and nonprimitive cells (centering translation) shown in Figure 2.2 makes up the 14 space lattices, called the Bravais lattices.[1] Of these 14, the face-centered cubic (FCC) and hexagonal space lattices are most relevant here. There is a higher classification hierarchy of crystalline materials based on the set of symmetry operators that can be applied to the 14 Bravais lattices in order to determine the point symmetry of the 3-D crystal. Classifying a crystalline structure as primitive or nonprimitive requires the determination of the lattice points within a unit cell; the number of lattice sites N in any given unit cell crystal structure may be determined by

$$N = N_{\text{Body}} + \frac{1}{2}N_{\text{Face}} + \frac{1}{4}N_{\text{Edge}} + \frac{1}{8}N_{\text{Corner}}, \tag{2.6}$$

where the subscript refers to the specific location of the site within the crystal structure.

A combination of symmetric operations such as rotation, reflection, and inversion that are compatible with translation give rise to 32 crystal classes (also known as point groups) into which crystalline structures can be further categorized. In addition, multiple combinations of these symmetric operators, which produce translations less than the magnitude of the basis vectors, give rise to 230 space groups. However, in our brief discussion of symmetry we must make two clear distinctions. First, a combination of symmetry elements whereby at least one atom or point of the crystal lattice remains fixed constitutes a point group. An example of this is the set of proper rotations about a screw axis. It should also be noted that the name *point group* derives from the idea that all symmetry elements performed on the crystal structure intersect at a single point. Second, the combination of symmetry elements involving translation that leaves no fixed point on the crystal lattice gives rise to space groups, that is, the set of glide planes. Although the focus of this book is not on the structure and classification of crystalline materials, these concepts serve as a fundamental background in the discussion of heteroepitaxy of semiconductors.

The wide variety of crystal structures among semiconductors presents both opportunities and challenges for the crystal grower. It is possible to create unique heterojunction devices and metastable structures by the proper choices of materials. Moreover, it is sometimes possible to determine the crystal structure of a particular epitaxial layer by the choice of substrate or growth conditions, adding another dimension to device design. On the other hand, it is also possible to end up with mixed phase material, which usually has degraded electronic properties. It is therefore the purpose of this section to describe the crystal structures exhibited by semiconductor materials of interest, as well as their characteristics and behavior relevant to heteroepitaxy.

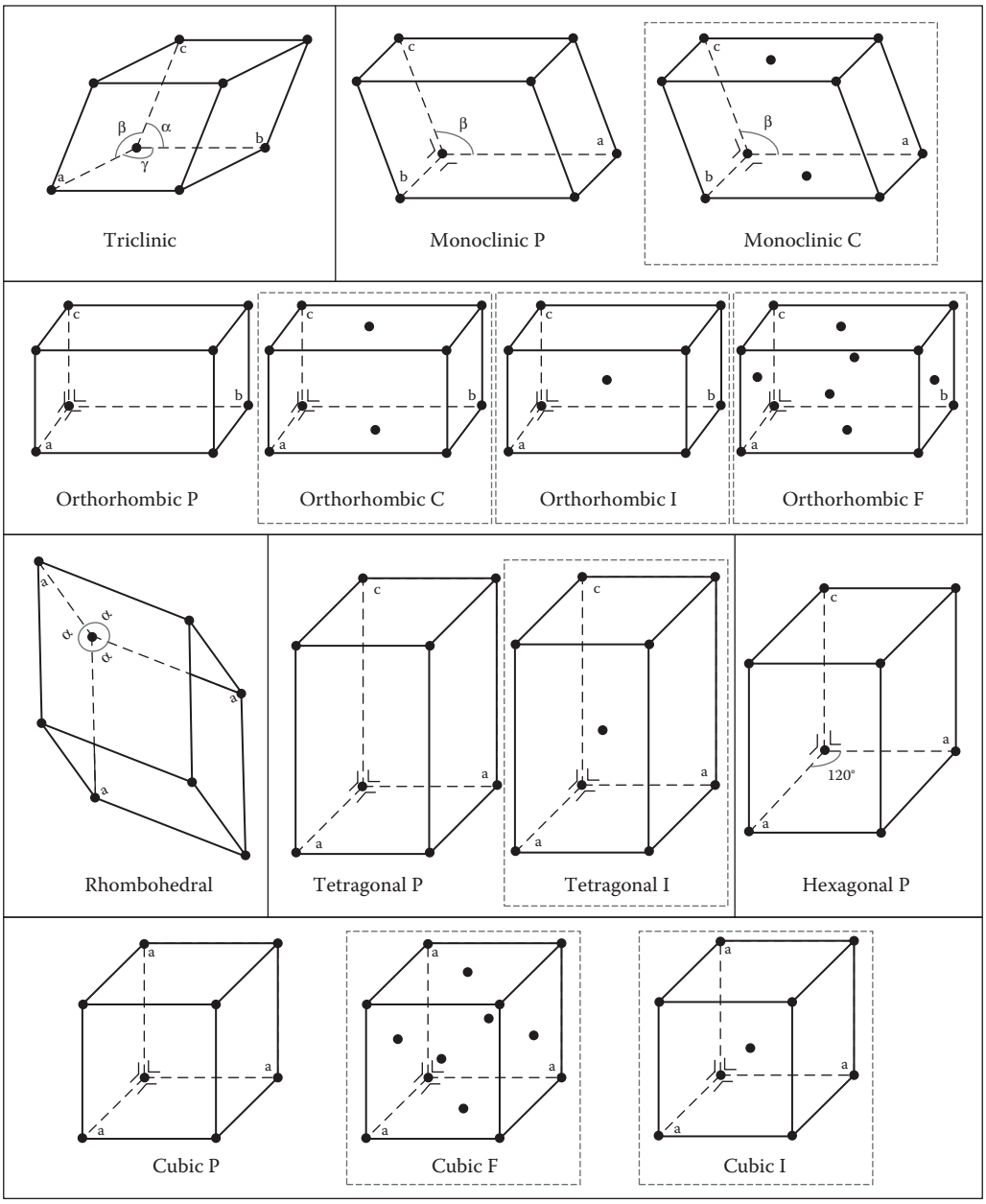


FIGURE 2.2
The 14 Bravais lattices. The cells within the dashed rectangles are the nonprimitive cells formed from centering translations. P indicates that the lattice structure is primitive, F indicates face centering, I indicates body centering, and C indicates base centering.

Some of most important crystallographic properties of semiconductors are the crystal structure and lattice constant(s). Also relevant to the growth of heteroepitaxial layers is the anisotropic behavior of the crystalline materials, especially the etching, nucleation, growth, and cleavage behavior. In many of the following subsections, materials will be lumped together in the cubic and hexagonal classes of crystals.

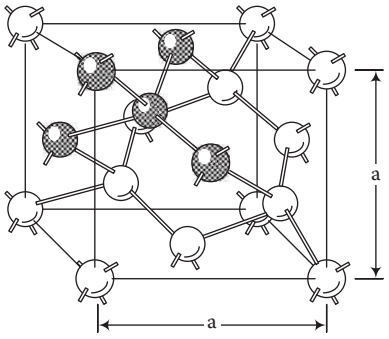


FIGURE 2.3
Diamond crystal structure. All atoms are of the same type (e.g., Si). The shaded spheres show an atom and its four nearest neighbors.

The technologically important semiconductors exhibit a number of different crystal structures. Silicon, germanium, and their alloys have the diamond structure. Many III-VI and II-VI semiconductors, including GaAs and InP, crystallize in the cubic zinc blende structure. GaN and related materials, as well as ZnS and other II-VI crystals, exhibit the hexagonal wurtzite structure. Some III-V and II-VI semiconductors can assume either a zinc blende or a wurtzite structure. SiC exhibits more than 250 different polytypes, including cubic, hexagonal, and rhombohedral variants.

2.2.1 Diamond Structure

The diamond structure is shared by Si, Ge, Si-Ge alloys, and α-Sn, as well as the diamond form of carbon. This structure belongs to the cubic class, with an FCC lattice and a basis of two atoms at each lattice point: one at the origin (0, 0, 0) and the other at a point with coordinates (¼*a*, ¼*a*, ¼*a*), where *a* is the lattice constant. Thus, the structure can be thought of as two interpenetrating FCC sublattices, one displaced from the other by one-quarter of the unit cell diagonal. The space group is $F\bar{4}3m$ (T_d^2).

Figure 2.3 shows the cubic unit cell of the diamond structure. The length of each side of the cubic unit cell is *a*, the lattice constant. The atoms are tetrahedrally bonded, and each atom in the structure is covalently bonded to its four nearest neighbors.

2.2.2 Zinc Blende Structure

A number of semiconductors exhibit the zinc blende* (ZB) structure, including GaAs, InP and other III-V semiconductors, CdTe, ZnSe, and other II-VI crystals, and also the cubic form of SiC. It is very similar to the diamond crystal structure, except that the two FCC sublattices are made up of two different types of atoms. As with the diamond structure, the space group is $F\bar{4}3m$ (T_d^2). The zinc blende structure is illustrated in Figure 2.4.

Because of the two types of atoms and therefore two compositionally different interpenetrating FCC sublattices, the zinc blende structure has a lower symmetry than the diamond structure. This can lead to interesting phenomena in the heteroepitaxy of zinc blende materials on diamond substrates. In (001) heteroepitaxy, it results in the nonequivalence

* The zinc blende structure is occasionally referred to as the "sphalerite" structure in the literature.

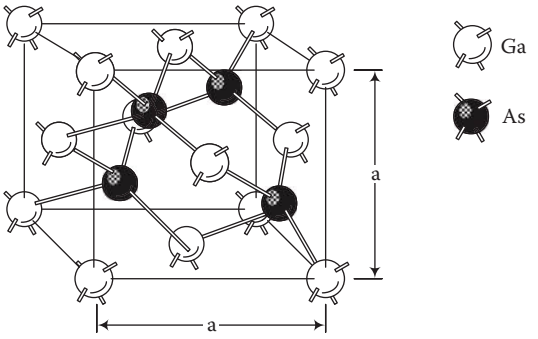


FIGURE 2.4
Zinc blende crystal structure. The white and black atoms belong to the two different sublattices (e.g., Ga and As).

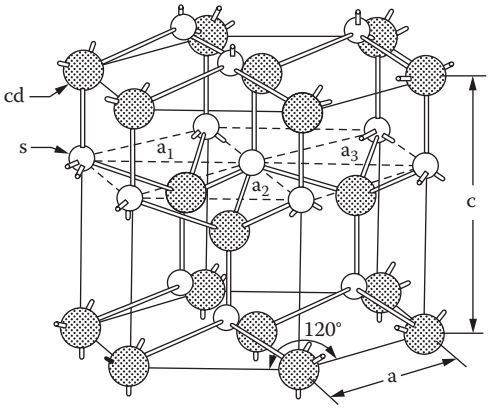


FIGURE 2.5
Wurtzite crystal structure. The white and shaded atoms belong to the two different sublattices.

of 60° dislocations along the [1$\bar{1}$0] and [110] directions, and these are referred to as α and β dislocations, respectively.

2.2.3 Wurtzite Structure

The wurtzite* (Wz) structure is common among III-nitrides such as AlN, GaN, and InN, and also some II-VI semiconductors. This structure comprises a hexagonal close-packed (HCP) lattice with a basis of two atoms; it can therefore be considered two interpenetrating HCP lattices. Because the unit cell is hexagonal, there are two lattice constants, *a* and *c*. (Here, *c* is the lattice constant in the direction parallel to the axis of sixfold rotational symmetry, as shown in Figure 2.5.) The two interpenetrating HCP lattices are made up of two different types of atoms, offset along the *c* axis by 5/8 of the cell height (5*c*/8). The space group is $P6_3mc$ (C_{6v}^4).

As with the zinc blende structure, the wurtzite crystal structure involves two types of atoms, A and B. Each atom A is bonded tetrahedrally to four nearest neighbors, which are type B. Because the nearest-neighbor configuration is the same as in the zinc blende

* The wurtzite structure is occasionally referred to as the "zincite" structure in the literature.

structure, the properties of the two structures are closely related if the second and third nearest-neighbor interactions are ignored.[2]

2.2.4 Silicon Carbide

SiC is unique among semiconductors in that it exists in many polytypes, the number of which has been reported to be as high as 250.[3] Many of these polytypes are hexagonal or rhombohedral, but there exists a cubic polytype as well. Each polytype is built up by stacking sheets of atoms. Each sheet can be represented as a close-packed two-dimensional (2-D) arrangement of spheres with sixfold symmetry, as shown in Figure 2.6. In the case of SiC, each close-packed sphere represents a silicon atom together with a carbon atom.

The polytypes differ in structure only in the stacking sequence for the bilayers of silicon and carbon atoms. This can be understood by referring to Figure 2.6. If the centers of the spheres in the first sheet (the basal plane for a hexagonal crystal) are at the positions marked A, the next sheet may be placed with the centers of the spheres over the points marked B or C. The various polytypes may be built up by stacking type A, B, and C layers in various sequences.

Regardless of the polytype, the atoms are tetrahedrally bonded with coordination number 4. That is, each silicon atom is bonded tetrahedrally with four carbon atoms, with a Si-C bond length of approximately 1.89 Å. Each carbon atom is similarly bonded to four silicon nearest neighbors. The tetrahedra have threefold symmetry, so for each stacking position A, B, or C, there are two variants in which the tetrahedra are rotated by 180° with respect to each other. The rotated variants are called A′, B′, and C′. Often the A, B, and C variants are referred to as "untwinned," whereas the A′, B′, and C′ variants are called "twinned."[4]

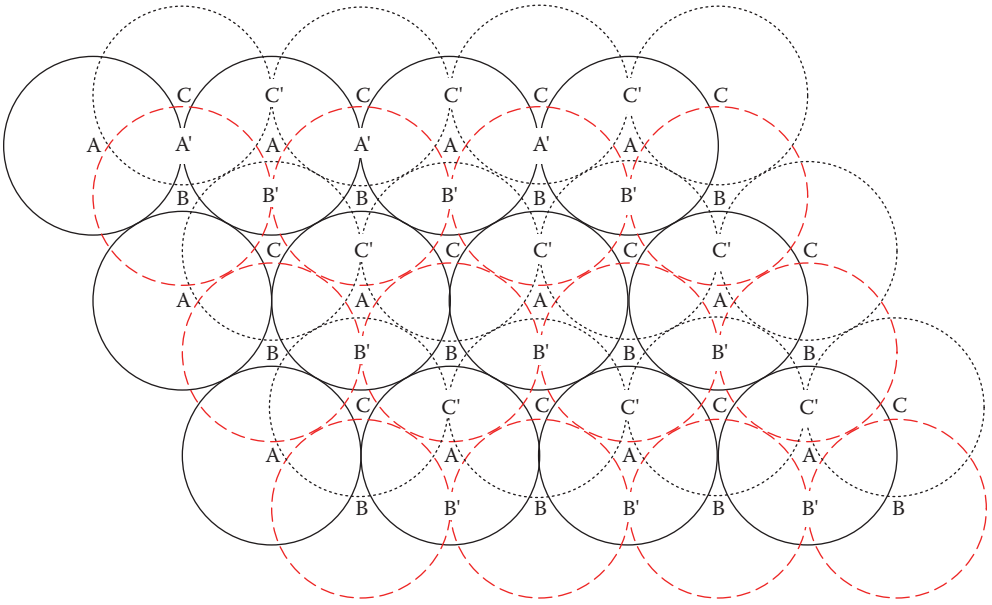


FIGURE 2.6
Close-packed spheres in two dimensions. The polytypes of SiC may be constructed by stacking such sheets of atoms, in which each sphere represents one silicon atom together with a carbon atom. The labels A, B, C, A′, B′, and C′ are described in the text.

Considering the stacking of the A, A′, B, B′, C, and C′ layers, only certain stacking sequences are allowed if corner sharing is to be maintained between tetrahedra.[5] Thus, an untwinned bilayer must be stacked on either an untwinned bilayer of the following letter (AB, BC, or CA) or a twinned bilayer of the preceding letter (AC′, BA′, or CB′). Similarly, a twinned bilayer must be stacked on either an untwinned bilayer of the following letter (A′B, B′C, or C′A) or a twinned bilayer of the preceding letter (A′C′, B′A′, or C′B′).

The resulting polytypes made by stacking bilayers in this way are cubic, hexagonal, or rhombohedral in structure. For example, the zinc blende polytype has the stacking sequence …ABCABC…. This polytype is referred to as 3C-SiC using the Ramsdell notation,[6] in which the C indicates a cubic structure and the 3 indicates the periodicity of the stacking sequence. Also of technological importance are the 4H and 6H polytypes. These are both hexagonal structures, with space group $P6_3mc$ (C_{6v}^4). The stacking sequences are …ABA′CABA′C… and …ABCB′A′C′ABCB′A′C′… for the 4H and 6H structures, respectively. Two other notable polytypes are 2H, which is a wurtzite structure, and 15R, which is rhombohedral.

Figure 2.7 shows the interstitial sites formed from the stacking of triangular tiling with atoms at all vertices in the close-packed cubic and hexagonal structures. The distance between vertices is twice the atomic radius of the constituent. In the cubic close-packed structure, the stacking of these tiles corresponds to the {111} planes, whereas in the HCP structure they make up the {001} planes. The interstices play a major role in the crystal structure because they may harbor other atoms and can alter the characteristic properties of the crystal. In the close-packed structure, if the triangular void has an atom directly above it, then the surrounding atoms form a tetrahedron, which is a structure with three ABA and one AAA triangular tile; this is commonly known as a tetrahedral interstice or a tetrahedral interstitial site. Figure 2.7a and b shows a tetrahedral arrangement of atoms and triangular tiles, respectively, whereas Figure 2.7e and g shows the tetrahedral interstitial in the cubic and hexagonal structures, respectively. In the octahedral interstice, the triangular void does not have an atom above it and the six atoms surrounding the interstitial site form an octahedron composed of one BBB, one AAA, three BAB, and three ABA triangular tiles. Figure 2.7c and d shows an octahedral arrangement of atoms and triangular tiles, respectively, whereas Figure 2.7f and h shows the octahedral interstitial in the cubic and hexagonal structures, respectively.

2.2.5 Miller Indices in Cubic Crystals

In cubic crystals such as the diamond and zinc blende structures, crystal planes and directions are denoted using Miller indices.

The Miller indices for a plane are obtained as follows. The intercepts of the plane with the three orthogonal axes $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are determined in terms of the lattice constant a; this yields three integers that may be positive or negative. The three smallest integers having the same ratios as the reciprocals of these intercepts are the Miller indices h, k, and l, and the plane is denoted (hkl). For example, consider the plane intercepting the $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ axes at ∞, ∞, a. The normalized intercepts are $\infty, \infty, 1$. Taking the reciprocals, we have $0, 0, 1$. These are integers so the plane is denoted (001). It is customary to indicate negative indices with an overbar rather than a minus sign. Thus, the plane (0–11) would usually be denoted as $(0\bar{1}1)$. Families of planes having the same symmetry are denoted by curly brackets, such as $\{hkl\}$. Therefore, the (100), $(\bar{1}00)$, (010), $(0\bar{1}0)$, (001), and $(00\bar{1})$ planes are collectively denoted as the {001} planes.

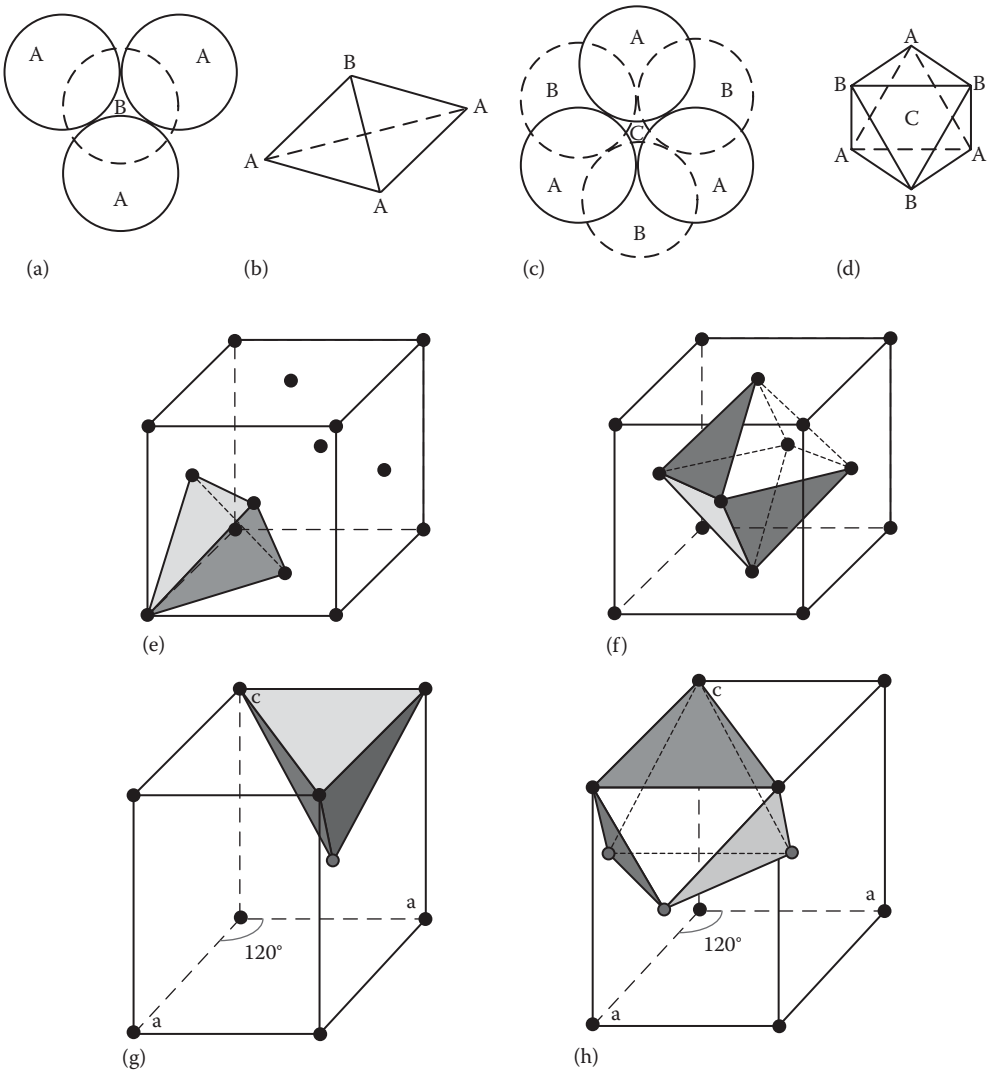


FIGURE 2.7
Atoms (a, c) and polyhedra (b, d) in (a, b) tetrahedral and (c, d) octahedral interstices. (e–h) Arrangement of atoms and polyhedra in the zinc blende and wurtzite configurations: tetrahedral (e, g) and octahedral (f, h) arrangements in the zinc blende structure, and tetrahedral (e, f) and octahedral (g, h) arrangements in the wurtzite structure.

Directions in a crystal are denoted by the smallest set of integers that have the same ratios as any vector in the direction. Thus, a direction is denoted [*uvw*] and a family of directions having the same symmetry is denoted ⟨*uvw*⟩. The [*uvw*] direction is perpendicular to the (*uvw*) plane. Several common planes and directions are illustrated in Figure 2.8.

2.2.6 Miller–Bravais Indices in Hexagonal Crystals

It is possible to use Miller indices, as defined above, for hexagonal crystals. Unfortunately, due to the lack of cubic symmetry, it is no longer true that the direction [*hkl*] is perpendicular

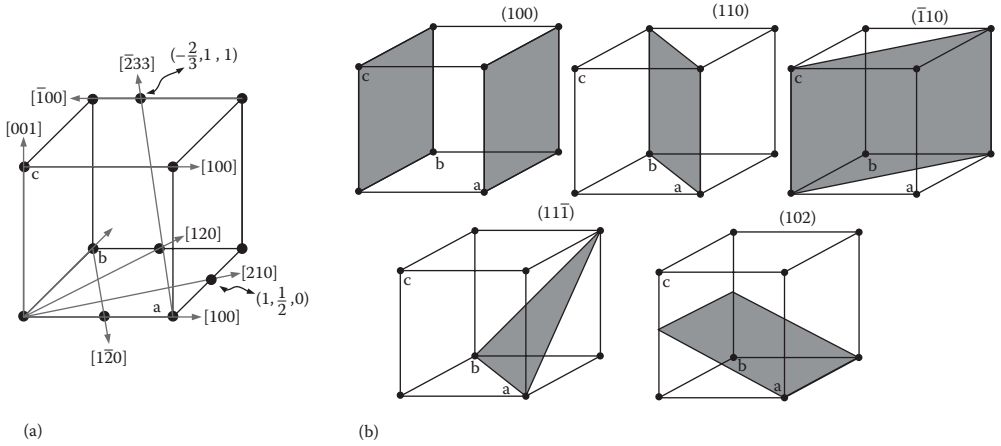


FIGURE 2.8
(a) Common directions and (b) planes in a cubic crystal system.

to the (*hkl*) plane in this case. Worse yet, planes with similar indices such as (100) and (001) do not have the same packing arrangement. For these reasons, the Miller–Bravais indices are used almost exclusively for hexagonal crystal structures.

The Miller–Bravais system uses four indices (*hkil*) to denote a plane. Here, the indices are obtained as in the cubic case, but the basal plane of the hexagonal unit cell is considered to have three axes, as shown in Figure 2.5. As a consequence, the index *i* is not independent but is given by

$$i = -(h + k) . \tag{2.6}$$

Despite the use of the superfluous index *i*, the Miller–Bravais system assigns similar indices to similar types of planes, which has led to its nearly universal adoption. Sometimes, the Miller–Bravais indices are abbreviated by omitting the superfluous index *i*, as in (*hkl*), especially when referring to x-ray reciprocal lattice points.

Referring to Figure 2.9, the *c* axis (axis of sixfold symmetry) is denoted [0001]. The basal plane is the (0001) plane. The a_1, a_2, and a_3 axes are denoted [2$\bar{1}\bar{1}$0], [$\bar{1}$2$\bar{1}$0], and [$\bar{1}\bar{1}$20], respectively.

Names have been given to several of the low-index planes in hexagonal crystals, as listed in Table 2.2. Examples of these planes are illustrated in Figure 2.10.

2.2.7 Computations and the Metric Tensor

In order to better understand the atomic arrangement in space and therefore the crystal structure, it is helpful to obtain relationships between distances and directions in a given crystalline material. In the following section, we will briefly introduce the mathematics behind coordinate transformation from the Cartesian reference frame to a particular system described by arbitrary basis vectors. In addition, we will introduce the metric tensor, a computation tool that may be used to determine the separation and angle between planes, the angle between a given plane and a particular direction, and the volume or density of the unit cell. It is therefore also useful in understanding the behavior of the crystal in relation to stress and strain, which are commonly encountered in heteroepitaxial growth.

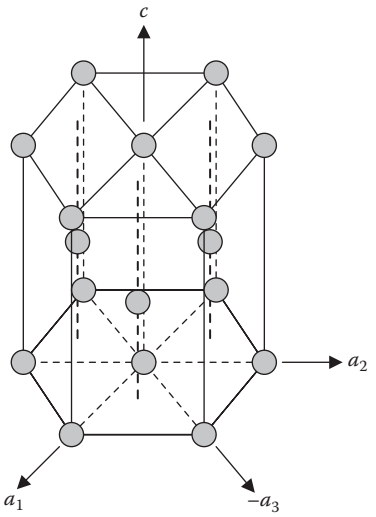


FIGURE 2.9
Axes for determination of the Miller–Bravais indices in a hexagonal unit cell.

TABLE 2.2
Planes of Hexagonal Semiconductor Crystals

Name	Example Planes in Miller–Bravais Notation
Basal plane	(0001)
Prism plane: first order	$(1\bar{1}00)$, $(\bar{1}100)$
Prism plane: second order	$(11\bar{2}0)$, $(\bar{2}110)$
Pyramidal plane: first order	$(10\bar{1}1)$, $(\bar{1}011)$
Pyramidal plane: second order	$(11\bar{2}2)$, $(\bar{1}\bar{1}22)$

2.2.7.1 Coordinate Transformation

In the Cartesian reference frame, we would usually find the distance D between two points located at the tips of the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ with coordinates $(x_1\hat{\mathbf{i}}, y_1\hat{\mathbf{j}}, z_1\hat{\mathbf{k}})$ and $(x_2\hat{\mathbf{i}}, y_2\hat{\mathbf{j}}, z_2\hat{\mathbf{k}})$ using the Pythagorean theorem so that

$$D = |\mathbf{r_1} - \mathbf{r_2}| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}\ . \tag{2.7}$$

However, in a nonorthogonal coordinate system with arbitrary basis vectors, such as the case of the triclinic crystal system, Equation 2.7 is not valid by itself. Coordinate transformation is necessary to express the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ in the Cartesian reference frame with basis vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ in terms of the alternate system with arbitrary basis vectors. Alternatively, we can make use of the vector dot product $\mathbf{a}\cdot\mathbf{b}$, also known as the scalar product, which is

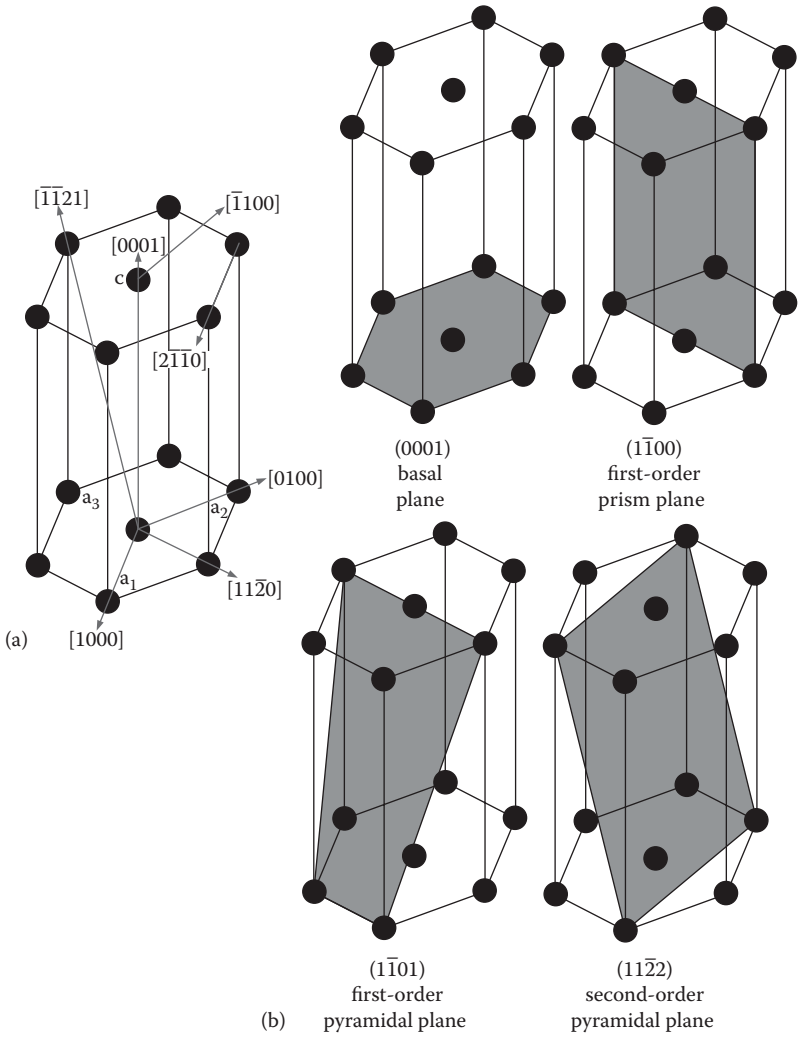


FIGURE 2.10
(a) Common directions and (b) planes of hexagonal semiconductor crystals.

defined as the projection of vector **a** onto the direction of **b** multiplied by the length of the vector **b**. In other words,

$$\mathbf{a} \cdot \mathbf{b} = |a|\,|b| \cos\gamma\,, \tag{2.8}$$

where $|a|$ and $|b|$ are the magnitudes (lengths) of vectors **a** and **b**, respectively, and γ is the angle between them. In addition, we can define the length of vector **a** as the square root of the dot product $\mathbf{a} \cdot \mathbf{a}$. From this definition, we can rewrite Equation 2.7 by accounting for the basis vectors of the Cartesian reference frame as

$$\begin{aligned} D &= |\,\mathbf{r}_1 - \mathbf{r}_2\,| = \sqrt{(\mathbf{r}_1 - \mathbf{r}_2)\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \\ &= \sqrt{(x_1 - x_2)^2 \hat{i}\cdot\hat{i} + (y_1 - y_2)^2 \hat{j}\cdot\hat{j} + (z_1 - z_2)^2 \hat{k}\cdot\hat{k}}\,. \end{aligned} \tag{2.9}$$

It can be seen that this expression reduces to Equation 2.7 in the case of the Cartesian reference frame because the dot product of parallel basis vectors is 1 and the dot product of orthogonal basis vectors is 0. Therefore, in the Cartesian reference frame $\hat{\mathbf{i}}\cdot\hat{\mathbf{i}}=\hat{\mathbf{j}}\cdot\hat{\mathbf{j}}=\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}=1$ and $\hat{\mathbf{i}}\cdot\hat{\mathbf{j}}=\hat{\mathbf{i}}\cdot\hat{\mathbf{k}}=\hat{\mathbf{j}}\cdot\hat{\mathbf{k}}=0$. A similar approach could be applied to any reference frame with nonorthogonal and arbitrary basis vectors.

2.2.7.2 The Metric Tensor

In Section 2.2.7.1, we saw how Equation 2.9 could be reduced to Equation 2.7 in the 3-D Cartesian reference frame through the use of the vector dot product. This section introduces the metric tensor, a computational tool that expresses the relation of Equation 2.9 in matrix form. We can therefore rewrite Equation 2.9 in matrix form as

$$D^2=\begin{bmatrix}x_1-x_2 & y_1-y_2 & z_1-z_2\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{i}}\cdot\hat{\boldsymbol{i}} & \hat{\boldsymbol{i}}\cdot\hat{\boldsymbol{j}} & \hat{\boldsymbol{i}}\cdot\hat{\boldsymbol{k}}\\ \hat{\boldsymbol{j}}\cdot\hat{\boldsymbol{i}} & \hat{\boldsymbol{j}}\cdot\hat{\boldsymbol{j}} & \hat{\boldsymbol{j}}\cdot\hat{\boldsymbol{k}}\\ \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{i}} & \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{j}} & \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}\end{bmatrix}\begin{bmatrix}x_1-x_2\\ y_1-y_2\\ z_1-z_2\end{bmatrix}. \tag{2.10}$$

The 3 × 3 matrix is known as the metric tensor, and it defines the geometrical characteristics of the Cartesian or any particular reference frame. In the case of a crystallographic reference frame (i.e., one of the seven fundamental crystals) where the crystal structure is defined by the lattice parameters $\{a,b,c,\alpha,\beta,\gamma\}$, the metric tensor is defined by

$$\mathbf{M}\equiv\begin{bmatrix}\mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c}\\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{c}\\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c}\end{bmatrix}=\begin{bmatrix}a^2 & ab\cos\gamma & ac\cos\beta\\ ba\cos\gamma & b^2 & bc\cos\alpha\\ ca\cos\beta & cb\cos\alpha & c^2\end{bmatrix}. \tag{2.11}$$

For a particular crystal system such as the ones given in Table 2.1, we can impose the appropriate restrictions on the angles or basis vectors and therefore obtain the appropriate metric tensor through the dot product of the basis vectors. The metric tensors for the seven fundamental crystal systems are given in Table 2.1. We can also rewrite Equation 2.10 for a generalized reference frame with a more compact form by

$$D^2=\sum_{i,j=1}^{n}\left(r_1-r_2\right)_i M_{ij}\left(r_1-r_2\right)_j, \tag{2.12}$$

where M_{ij} is the metric tensor, n is the dimensionality of the vector space, and the indices i and j are used to specify the row and column of the vector space, respectively. It should also be noted that the components of the matrix to the left of the metric tensor are written as a row, whereas those of the one on the right are written in a column. The metric tensor could be used to describe the dot product between two vectors. Following the relation described in Equation 2.9 and the generalized expression of Equation 2.12, we can write the dot product between vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ as

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = \sum_{i,j=1}^{n} r_i M_{ij} r_j \ . \tag{2.13}$$

For example, we can determine the dot product of the vectors $\mathbf{r}_1 = [uvw]$ and $\mathbf{r}_2 = [mno]$ using the triclinic metric tensor as

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ba\cos\gamma & b^2 & bc\cos\alpha \\ ca\cos\beta & cb\cos\alpha & c^2 \end{bmatrix} \begin{bmatrix} m \\ n \\ o \end{bmatrix}$$

$$= \begin{pmatrix} mua^2 + nvb^2 + owc^2 + ab(nu + mv)\cos\gamma \\ +ac(ou + mw)\cos\beta + bc(ov + nw)\cos\alpha \end{pmatrix}. \tag{2.14}$$

In addition, the length of a vector $\mathbf{r}_1 = [uvw]$ can be determined using the general triclinic metric tensor as

$$|\mathbf{r}_1| = \sqrt{\mathbf{r}_1 \cdot \mathbf{r}_1} = \sqrt{\begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ba\cos\gamma & b^2 & bc\cos\alpha \\ ca\cos\beta & cb\cos\alpha & c^2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}}$$

$$= \sqrt{\left(u^2a^2 + v^2b^2 + w^2c^2 + 2abvu\cos\gamma + 2acuw\cos\beta + 2bcvw\cos\alpha\right)}. \tag{2.15}$$

2.2.7.3 Distances between Lattice Points and Lengths of Vectors

Following the analysis in Section 2.2.7.2, we can use the metric tensor for various crystallographic computations. Using Equation 2.12, we can find the distance between two atoms with coordinates (u, v, w) and (m, n, o) in an arbitrary reference frame described by lattice parameters $\{a, b, c, \alpha, \beta, \gamma\}$ as follows:

$$D^2 = \begin{bmatrix} u-m & v-n & o-w \end{bmatrix} \begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ba\cos\gamma & b^2 & bc\cos\alpha \\ ca\cos\beta & cb\cos\alpha & c^2 \end{bmatrix} \begin{bmatrix} u-m \\ v-n \\ w-o \end{bmatrix}$$

$$= \begin{pmatrix} a^2(m-u)^2 + b^2(n-v)^2 + c^2(o-w)^2 + 2ab(m-u)(n-v)\cos\gamma \\ +2ac(o-w)(m-u)\cos\beta + 2bc(o-w)(n-v)\cos\alpha \end{pmatrix}. \tag{2.16}$$

In the tetragonally distorted cubic lattice (or the tetragonal crystal system) the distance between two atoms with coordinates (u, v, w) and (m, n, o) reduces to

$$D = \sqrt{a^2\left((m-u)^2 + (n-v)^2\right) + c^2(o-w)^2} \ . \tag{2.17}$$

The approach described above could also be applied to find the length l of a particular vector [uvw] and is given by

$$l_{[uvw]} = \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ba\cos\gamma & b^2 & bc\cos\alpha \\ ca\cos\beta & cb\cos\alpha & c^2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

$$= \begin{pmatrix} a^2u^2 + b^2v^2 + c^2w^2 + 2abuv\cos\gamma \\ +2acwu\cos\beta + 2bcwv\cos\alpha \end{pmatrix}. \tag{2.18}$$

In the tetragonal system, the length l of a vector [uvw] is given by

$$l_{[uvw]} = \sqrt{a^2\left(u^2 + v^2\right) + c^2w^2}\ . \tag{2.19}$$

In the case of an undistorted cubic structure, Equation 2.18 reduces to

$$l_{[uvw]} = \sqrt{a^2\left(u^2 + v^2 + w^2\right)}\ . \tag{2.20}$$

When dealing with the hexagonal crystal system, we can apply a similar approach to determine the length of a vector. However, if we wanted to show calculations that include the degenerate basis vector, the metric tensor could be modified as follows:

$$\mathbf{M}_{Hex}^{4\times4} = \begin{bmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot(\mathbf{a}+\mathbf{b}) & \mathbf{a}\cdot\mathbf{c} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot(\mathbf{a}+\mathbf{b}) & \mathbf{b}\cdot\mathbf{c} \\ (\mathbf{a}+\mathbf{b})\cdot\mathbf{a} & (\mathbf{a}+\mathbf{b})\cdot\mathbf{b} & (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) & (\mathbf{a}+\mathbf{b})\cdot\mathbf{c} \\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot(\mathbf{a}+\mathbf{b}) & \mathbf{c}\cdot\mathbf{c} \end{bmatrix}$$

$$= \begin{bmatrix} a^2 & -\frac{a^2}{2} & \frac{a^2}{2} & 0 \\ -\frac{a^2}{2} & a^2 & \frac{a^2}{2} & 0 \\ \frac{a^2}{2} & \frac{a^2}{2} & a^2 & 0 \\ 0 & 0 & 0 & c^2 \end{bmatrix}. \tag{2.21}$$

The above 4 × 4 metric tensor for the hexagonal crystal system is equivalent to the 3 × 3 matrix given in Table 2.1, and it can be shown that the following calculations will yield the same result. Thus,

$$\begin{bmatrix} h & k & l \end{bmatrix} \begin{bmatrix} a^2 & -\frac{a^2}{2} & 0 \\ -\frac{a^2}{2} & a^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix} \begin{bmatrix} h \\ k \\ l \end{bmatrix} \tag{2.22}$$

and

$$\begin{bmatrix} h & k & -(h+k) & l \end{bmatrix} \begin{bmatrix} a^2 & -\frac{a^2}{2} & \frac{a^2}{2} & 0 \\ -\frac{a^2}{2} & a^2 & \frac{a^2}{2} & 0 \\ \frac{a^2}{2} & \frac{a^2}{2} & a^2 & 0 \\ 0 & 0 & 0 & c^2 \end{bmatrix} \begin{bmatrix} h \\ k \\ -(h+k) \\ l \end{bmatrix} \tag{2.23}$$

are equivalent. The 4×4 matrix may be helpful in the discussion of distortion in hexagonal crystal structures.

2.2.7.4 Angle between Vector Directions

In addition to calculating distances between lattice points, we can use the previously described formulations to determine the angle between two vectors using the dot product and recognizing that each crystal direction may be specified by a translation vector. Two vectors [*uvw*] and [*mno*] can be written in terms of their respective translation vectors as $\mathbf{t}_{[uvw]}$ and $\mathbf{t}_{[mno]}$ and the angle between them is given by

$$\theta = \cos^{-1}\left(\frac{\mathbf{t}_{[uvw]} \cdot \mathbf{t}_{[mno]}}{|\mathbf{t}_{[uvw]}|\,|\mathbf{t}_{[mno]}|} \right), \tag{2.24}$$

where $\mathbf{t}_{[\ldots]}$ and $|\mathbf{t}_{[\ldots]}|$ can be found by methods described in Equations 2.14 and 2.15, respectively. Combining the results described in Equations 2.14 and 2.15, we can determine the angle between two vectors in any arbitrary reference frame. In the case of the tetragonal crystal system, the angle between two vectors $\mathbf{r}_1 = [uvw]$ and $\mathbf{r}_2 = [mno]$ is given by

$$\theta = \cos^{-1}\left(\frac{a^2(um+vn)+c^2wo}{\sqrt{a^2(u^2+v^2)+c^2w^2}\sqrt{a^2(m^2+n^2)+c^2o^2}} \right). \tag{2.25}$$

In the undistorted cubic lattice the expression given in Equation 2.24 reduces to

$$\theta = \cos^{-1}\left(\frac{(um+vn+wo)}{\sqrt{(u^2+v^2+w^2)}\sqrt{(m^2+n^2+o^2)}} \right). \tag{2.26}$$

In the hexagonal crystal system, the angle between vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ is given by

$$\theta = \cos^{-1}\left(\frac{a^2\left(2um - un - vm + 2vn\right) + 2c^2wo}{2\sqrt{a^2\left(u^2 - uv + v^2\right) + c^2w^2}\sqrt{a^2\left(m^2 - nm + n^2\right) + c^2o^2}} \right). \tag{2.27}$$

2.2.7.5 Volume of a Unit Cell

The volume of a unit cell (V_C) described by the basis vectors **a**, **b**, and **c** may be calculated using the triple scalar product of the basis vector as follows:

$$V_C = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \,. \tag{2.28}$$

In the most general case, the volume of the triclinic unit cell is given by

$$V_C = abc\sqrt{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma} \,. \tag{2.29}$$

It also follows that the square root of the determinant of the metric tensor yields the volume of the unit cell. Therefore,

$$V_C = \sqrt{\det(M_{ij})} \,. \tag{2.30}$$

The vector cross product is often used to define crystal planes by their normal, and its concept may be extended to the formation of the reciprocal space lattice.

2.2.7.6 Reciprocal Basis Vectors and Reciprocal Metric Tensor

The reciprocal lattice and reciprocal lattice vectors are useful in consideration of diffraction by crystals. The set of reciprocal lattice vectors $\mathbf{r}_j^*$ is defined by

$$\mathbf{r}_i \cdot \mathbf{r}_j^* = \delta_{ij} = \begin{cases} 1 & if\ i = j \\ 0 & if\ i \neq j \end{cases}, \tag{2.31}$$

where the asterisk (*) indicates the use of a reciprocal basis vector and δ_{ij} is the Kronecker delta, also known as the identity matrix. In other words, for a 3-D lattice, Equation 2.30 may be written as

$$\begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{r}_1^* & \mathbf{r}_1 \cdot \mathbf{r}_2^* & \mathbf{r}_1 \cdot \mathbf{r}_3^* \\ \mathbf{r}_2 \cdot \mathbf{r}_1^* & \mathbf{r}_2 \cdot \mathbf{r}_2^* & \mathbf{r}_2 \cdot \mathbf{r}_3^* \\ \mathbf{r}_3 \cdot \mathbf{r}_1^* & \mathbf{r}_3 \cdot \mathbf{r}_2^* & \mathbf{r}_3 \cdot \mathbf{r}_3^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.32}$$

The definition of Equation 2.31 indicates that vector $\mathbf{r}_1^*$ is perpendicular to vectors $\mathbf{r}_2$ and $\mathbf{r}_3$ since the vector dot product is zero. Similarly, the vector $\mathbf{r}_2^*$ is perpendicular to vectors $\mathbf{r}_1$ and $\mathbf{r}_3$ and vector $\mathbf{r}_3^*$ is perpendicular to vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. In addition, if a vector is normal

(or orthogonal) to two vectors, then it must also be parallel to the cross product of these two vectors. Because $\mathbf{r}_1 \cdot \mathbf{r}_1^* = 1$, the reciprocal basis vectors may be determined by

$$\mathbf{a}^* = \frac{\mathbf{b} \times \mathbf{c}}{V_C}, \tag{2.33}$$

$$\mathbf{b}^* = \frac{\mathbf{c} \times \mathbf{a}}{V_C}, \tag{2.34}$$

and

$$\mathbf{c}^* = \frac{\mathbf{a} \times \mathbf{b}}{V_C}, \tag{2.35}$$

where $V_C = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ is the volume of the unit cell described by the basis vectors **a**, **b**, and **c**. Although a particular lattice could be described in many ways, the results of Equation 2.31 impose special requirements on the selection of the reciprocal basis vectors. Therefore, a vector quantity **r** in a particular reference frame and its reciprocal must obey the following relation:

$$\mathbf{r} = \sum_{i=1}^{N} p_i a_i = \sum_{j=1}^{N} p_j^* a_j^* \ . \tag{2.36}$$

It can be seen from the above relation that we can express the components of a vector **r** in terms of the reciprocal basis vectors in a fashion similar to the one using direct basis vectors. This analysis enables us to construct the reciprocal metric tensor defined by

$$\mathbf{M}^* \equiv \begin{bmatrix} \mathbf{a}^* \cdot \mathbf{a}^* & \mathbf{a}^* \cdot \mathbf{b}^* & \mathbf{a}^* \cdot \mathbf{c}^* \\ \mathbf{b}^* \cdot \mathbf{a}^* & \mathbf{b}^* \cdot \mathbf{b}^* & \mathbf{b}^* \cdot \mathbf{c}^* \\ \mathbf{c}^* \cdot \mathbf{a}^* & \mathbf{c}^* \cdot \mathbf{b}^* & \mathbf{c}^* \cdot \mathbf{c}^* \end{bmatrix} = \begin{bmatrix} a^{*2} & a^* b^* \cos\gamma^* & a^* c^* \cos\beta^* \\ b^* a^* \cos\gamma^* & b^{*2} & b^* c^* \cos\alpha^* \\ c^* a^* \cos\beta & c^* b^* \cos\alpha^* & c^{*2} \end{bmatrix}, \tag{2.37}$$

where $\{a^*, b^*, c^*, \alpha^*, \beta^*, \gamma^*\}$ are the parameters of the reciprocal lattice. All the mathematics described in the previous sections apply to the reciprocal space lattice. However, the reciprocal metric tensor is necessary in the calculation of distances and angles between planes. The reciprocal metric tensors for the tetragonal, hexagonal, and tetragonal crystal systems are given in Table 2.3. In addition, there is a simple relation between the metric tensor and the metric tensor for the reciprocal lattice given by

$$\mathbf{M} = \left(\mathbf{M}^*\right)^{-1}. \tag{2.38}$$

A directional (translation) vector in the reciprocal space is usually denoted using the symbol **g** and its components are given using the Miller indices as (*hkl*). Therefore, the reciprocal lattice vector $\mathbf{g}_{[hkl]}$ may be written as

$$\mathbf{g} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \ . \tag{2.39}$$

TABLE 2.3

Reciprocal Metric Tensor for the Tetragonal, Hexagonal, and Cubic Crystal Systems

Crystal System {Lattice Parameters} Requirements	Reciprocal Metric Tensor
Tetragonal $\{a, a, c, 90°, 90°, 90°\}$ $a = b$ $\alpha = \beta = \gamma = 90°$	$\begin{bmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{c^2} \end{bmatrix}$
Hexagonal $\{a, a, c, 90°, 90°, 120°\}$ $a = b$ $\alpha = \beta = 90°$ $\gamma = 120°$	$\begin{bmatrix} \frac{4}{3a^2} & \frac{2}{3a^2} & 0 \\ \frac{2}{3a^2} & \frac{4}{3a^2} & 0 \\ 0 & 0 & \frac{1}{c^2} \end{bmatrix}$
Cubic $\{a, a, a, 90°, 90°, 90°\} a = b = c$ $\alpha = \beta = \gamma = 90°$	$\begin{bmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{a^2} \end{bmatrix}$

2.2.7.7 Distances and Angles Involving Planes

From the description of Miller indices, it becomes apparent that a reciprocal lattice vector $\mathbf{g}_{[hkl]}$ is perpendicular to the plane (hkl). Therefore, the unit normal $\mathbf{n}_{[hkl]}$ is given by

$$\mathbf{n}_{[hkl]} = \frac{\mathbf{g}_{[hkl]}}{|\mathbf{g}_{[hkl]}|}. \tag{2.40}$$

In addition, the interplanar spacing d_{hkl} may be calculated using the length of the reciprocal lattice vector $\mathbf{g}_{[hkl]}$ by the following relation:

$$d_{hkl} = \frac{1}{|\mathbf{g}_{[hkl]}|}. \tag{2.41}$$

Using an approach similar to that in Section 2.2.7.3, the distance between $\{hkl\}$ planes in the tetragonal crystal system is given by

$$d_{\{hkl\}} = \frac{1}{|\mathbf{g}_{[hkl]}|} = \frac{1}{\sqrt{\mathbf{g}_{[hkl]} \cdot \mathbf{g}_{[hkl]}}}$$

$$= \left[\begin{bmatrix} h & k & l \end{bmatrix} \begin{bmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{c^2} \end{bmatrix} \begin{bmatrix} h \\ k \\ l \end{bmatrix} \right]^{-\frac{1}{2}}$$

$$= \left[\frac{(h^2 + k^2)}{a^2} + \frac{l^2}{c^2} \right]^{-\frac{1}{2}}. \tag{2.42}$$

In the case of the undistorted cubic lattice, the expression given in Equation 2.41 reduces to

$$d_{\{hkl\}} == \frac{a}{\sqrt{h^2 + k^2 + l^2}}. \tag{2.43}$$

Similarly, for the hexagonal system the distance between {*hkil*} planes may be determined by

$$d_{\{hkl\}} = \left[\begin{bmatrix} h & k & l \end{bmatrix} \begin{bmatrix} \frac{4}{3a^2} & \frac{2}{3a^2} & 0 \\ \frac{2}{3a^2} & \frac{4}{3a^2} & 0 \\ 0 & 0 & \frac{1}{c^2} \end{bmatrix} \begin{bmatrix} h \\ k \\ l \end{bmatrix} \right]^{-\frac{1}{2}}$$

$$= \left[\frac{4(h^2 + hk + k^2)}{3a^2} + \frac{l^2}{c^2} \right]^{-\frac{1}{2}}. \tag{2.44}$$

If we wanted to use the 4×4 reciprocal metric tensor including the superfluous index $i = -(h+k)$ in the calculations for the hexagonal crystal system, Equation 2.44 would be modified to

$$d_{\{hkil\}} = \left[\begin{bmatrix} h & k & -(h+k) & l \end{bmatrix} \begin{bmatrix} \frac{4}{3a^2} & \frac{2}{3a^2} & \frac{2}{3a^2} & 0 \\ \frac{2}{3a^2} & \frac{4}{3a^2} & \frac{2}{3a^2} & 0 \\ \frac{2}{3a^2} & \frac{2}{3a^2} & \frac{4}{3a^2} & 0 \\ 0 & 0 & 0 & \frac{1}{c^2} \end{bmatrix} \begin{bmatrix} h \\ k \\ -(h+k) \\ l \end{bmatrix} \right]^{-\frac{1}{2}}$$

$$= \left[\frac{4(h^2 + hk + k^2)}{3a^2} + \frac{l^2}{c^2} \right]^{-\frac{1}{2}}. \tag{2.45}$$

Once again, it can be seen that the results of Equations 2.44 and 2.45 yield the same value for the distance between {*hkil*} planes in the hexagonal crystal system.

The angle between two planes may be computed using their respective normal vectors. Similar to the theory explained in Section 2.2.7.4, the angle ϕ between two planes (*uvw*) and (*mno*) with normal vectors $\mathbf{g}_{[uvw]}$ and $\mathbf{g}_{[mno]}$, respectively, is given by

$$\phi = \cos^{-1}\left(\frac{\mathbf{g}_{[uvw]} \cdot \mathbf{g}_{[mno]}}{|\mathbf{g}_{[uvw]}|\,|\mathbf{g}_{[mno]}|}\right)$$

$$= \cos^{-1}\left(\frac{\begin{bmatrix} u & v & w \end{bmatrix}\begin{bmatrix}\mathbf{M}^*\end{bmatrix}\begin{bmatrix} m \\ n \\ o \end{bmatrix}}{\sqrt{\begin{bmatrix} u & v & w \end{bmatrix}\begin{bmatrix}\mathbf{M}^*\end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix}}\sqrt{\begin{bmatrix} m & n & o \end{bmatrix}\begin{bmatrix}\mathbf{M}^*\end{bmatrix}\begin{bmatrix} m \\ n \\ o \end{bmatrix}}}\right). \tag{2.46}$$

Equation 2.45 is identical in nature to Equation 2.24, with the only exception being the use of the reciprocal metric tensor. Thus, in the tetragonal system the angle between two planes (*uvw*) and (*mno*) may be computed by

$$\phi = \cos^{-1}\left(\frac{a^2wo + c^2(um + vn)}{a^2c^2\sqrt{\frac{(u^2+v^2)}{a^2} + \frac{w^2}{c^2}}\sqrt{\frac{(m^2+n^2)}{a^2} + \frac{o^2}{c^2}}}\right). \tag{2.47}$$

For the cubic crystal system the expression of Equation 2.46 reduces to

$$\phi = \cos^{-1}\left(\frac{(wo + um + vn)}{\sqrt{u^2+v^2+w^2}\sqrt{m^2+n^2+o^2}}\right). \tag{2.48}$$

In the case of the hexagonal system, the angle between two planes (*hkil*) and (*mnpo*) is

$$\phi = \cos^{-1}\left(\frac{3a^2lo + 2c^2\left(h(2m+n+p) + k(m+2n+p) + i(m+n+2p)\right)}{a^2c^2\left[\sqrt{\frac{4\left(h^2+i^2+ik+k^2+h(i+k)\right)}{a^2} + \frac{3l^2}{c^2}} * \sqrt{\frac{3a^2o^2 + 4c^2\left(m^2+n^2+np+p^2+m(n+p)\right)}{a^2c^2}}\right]}\right). \tag{2.49}$$

Recognizing that $i = -(h+k)$ and $p = -(m+n)$, the expression in Equation 2.49 reduces to

$$\phi = \cos^{-1}\left(\frac{3a^2lo + 2c^2\left(h(2m+n) + k(m+2n)\right)}{a^2c^2\left[\sqrt{\frac{4\left(h^2+hk+k^2\right)}{a^2} + \frac{3l^2}{c^2}}\sqrt{\frac{4\left(m^2+mn+n^2\right)}{a^2c^2} + \frac{3o^2}{c^2}}\right]}\right). \tag{2.50}$$

In the cases where the angle between the planes (*hkl*) and the crystal direction [*mno*] needs to be computed, we can use Equation 2.46, with the only exception that the left side of the equation is modified to $90 - \phi$.

2.2.8 Orientation Effects

Because the semiconductor materials used in devices are single-crystal materials, many of their properties and fabrication processes are different for the various crystal faces. Thus, the etching and cleaning, epitaxial growth, cleavage, electronic properties, and defect structure are orientation dependent. Some of these orientation effects will be described here.

2.2.8.1 Diamond Semiconductors

A number of the orientation effects arising in a crystal may be understood based on the packing densities, interplanar spacings, or bond densities of the low-index planes. In diamond structure materials such as silicon, the atom densities on the principal planes are in the ratios $\{100\} : \{110\} : \{111\} = 1 : 1.414 : 1.155$. Because atoms in these planes have two, one, and one dangling bonds to the next plane, respectively, the densities of interplanar bonds are in the ratios $\{100\} : \{110\} : \{111\} = 1 : 0.707 : 0.577$.

The {111} planes generally exhibit the lowest growth rates and the lowest etch rates. This is because a {111} surface will have the lowest density of dangling (interplanar) bonds (but a high density of intraplanar bonds). Thus, epitaxial growth under kinetically controlled conditions will tend to reveal the slow-growth {111} planes and cause faceting. Similarly, surface-sensitive etches (crystallographic etches) will tend to delineate the {111} planes.

The natural cleavage planes of silicon are the {111} planes, due to their weaker interplanar bonding. In the case of Si (001), the {111} cleavage planes meet the surface at an angle of 54.74° along the ⟨110⟩ directions. Therefore, the crystal will naturally break apart into rectangular dice if scribed along the ⟨110⟩ directions, which meet at 90° angles on the surface.

Cleavage of an Si {111} wafer is complicated by the fact that the three other {111} planes meet the surface along ⟨110⟩ directions, which are mutually at 60° to each other. Therefore, if rectangular dice are to be cut, only two sides of the rectangle may follow natural cleavage planes. The other two sides will have jagged edges, the teeth of which will be made up of the natural cleavage planes. The cleavage planes meet the (111) surface at an angle of 54.74°, as shown in Figure 2.11.

2.2.8.2 Zinc Blende Semiconductors

The orientation effects in crystals such as GaAs are different from those in silicon because of the polar nature of the zinc blende structure. Whereas the bonding in Si or Ge is entirely covalent, GaAs has a partly ionic character due to the different atoms on the two FCC sublattices. The Ga and As ions in the crystal (cations and anions, respectively) take on net positive and negative electrical charges, respectively. The other zinc blende crystals behave in a similar fashion, but GaAs will be discussed for the purpose of specificity.

Owing to this polar nature, the <111> directions are not equivalent in a zinc blende crystal such as GaAs. In the [111] direction, the crystal may be built up by stacking alternating layers of Ga and As atoms. However, these layers are not equally spaced but are stacked as …Ga-As—Ga-As—Ga-As…. Each Ga atom will be tetrahedrally bonded to one As atom in the layer directly below, and to three As atoms in the layer directly above. On the other

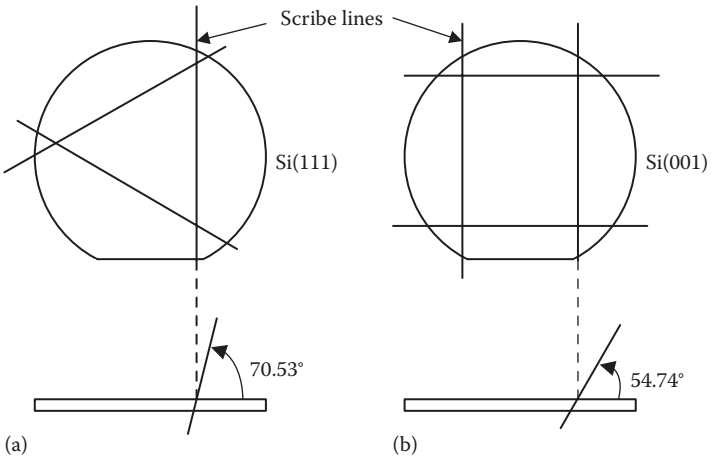


FIGURE 2.11
Cleavage planes of (111) (a) and (001) (b) Si wafers.

hand, each As atom will be tetrahedrally bonded to three Ga atoms in the layer directly below and one Ga atom above. This means that an As atom on the (111) surface will have one dangling bond, whereas a Ga atom on the (111) surface will have three dangling bonds. For this reason, this surface will be made up entirely of As atoms and is called the (111)As face. By the same token, the $(\bar{1}\,\bar{1}\,\bar{1})$ face will comprise only Ga atoms and is called the (111) Ga face.

The (111)As face is electronically more active than the (111)Ga face. This is because on the (111)As face, pentavalent As atoms are bonded to three Ga atoms in the underlying layer, leaving two free electrons each. In the (111)Ga face, however, the trivalent Ga atoms each participate in bonding with three As atoms from the layer below, leaving no free electrons. As a consequence, crystal growth and etching both occur rapidly on the (111)As face but slowly on the (111)Ga face. Thus, the (111)As face can be polished to a mirror finish and allows smooth epitaxial layers with good crystal quality, whereas the (111)Ga face is difficult to polish and epitaxial layers on this orientation tend to have poor morphology.

GaAs does not cleave at {111} planes because of their coulombic attraction. (These planes are alternating layers of Ga and As atoms, with net positive and negative charges, respectively.) Instead, GaAs cleaves on the neutral {110} planes. This is of practical importance because a GaAs (001) wafer may be cleaved into rectangular dice with vertical sides. In this case, the edges of the cleaved die will have (110), $(\bar{1}10)$, $(1\bar{1}0)$, and $(\bar{1}\bar{1}0)$ sides. Their intersections with the surface are along the $[1\bar{1}0]$, [110], $[\bar{1}10]$, and [110] directions, which are mutually at right angles. The cleavage behavior is useful in the packaging of devices, and also makes it possible to cleave a laser diode (LD) cavity with perfectly parallel end mirrors.

There are no remarkable differences in growth rates or etching rates among the low-index planes of GaAs or similar materials. There is, in fact, no good selective (anisotropic) etch for GaAs. However, the epitaxial growth and etching rates are often in the order (011) > (111)As > (001) > (111)Ga.

2.2.8.3 Wurtzite Semiconductors

Like the zinc blende materials, wurtzite semiconductors such as GaN are polar. As such, the [0001] and $[000\bar{1}]$ directions are not equivalent. In the [0001] direction, the crystal may

be built up by stacking alternating close-packed layers of Ga and N atoms with unequal spacings, as …Ga-N—Ga-N—Ga-N…. In the [0001] direction, each Ga atom is bonded to one N atom in the layer below and three N atoms in the layer above, but each N atom is bonded to three Ga atoms below and one Ga atom above. Based on this, a N atom on a (0001) surface will have one dangling bond, but a Ga atom on the same surface would have three dangling bonds. For this reason, the (0001) face is made up entirely of N atoms and is called the (0001)N face. Following the same arguments, the $(000\bar{1})$ face is made up entirely of Ga atoms and is called the (0001)Ga face.

The (0001)N face is more electronically active than the (0001)Ga face. This is because the pentavalent N atoms on the surface of the (0001)N face have three electrons bonded to Ga atoms in the layer below, but the other two valence electrons are free. On the (0001)Ga face, all three valence electrons from each Ga atom are bonded with As atoms in the underlying layer, leaving zero free electrons.

The natural cleavage planes of wurtzite crystals like GaN are the first-order prism planes of type $\{1\bar{1}00\}$. In the case of GaN (0001), these cleavage planes are perpendicular to the (*c*-face) surface, and intersect the surface along $\langle 11\bar{2}0\rangle$ directions, which are mutually at 60° angles.

Up to the present time, III-nitride devices have been fabricated almost exclusively on dissimilar substrates, so the cleavage behavior of these substrates is also important. The most commonly used substrate for III-nitride heteroepitaxy is sapphire (α-Al_2O_3). For sapphire, the natural cleavage planes are $\{1\bar{1}02\}$ planes (the so-called *R*-faces). For this reason, GaN (0001) is sometimes grown heteroepitaxially on Al_2O_3 $(11\bar{2}0)$, "*a*-face sapphire," so that the natural cleavage planes of the GaN and sapphire line up approximately.[7–9]

The introduction of semipolar and nonpolar orientations has positively changed the landscape for nitride-based light-emitting diodes (LEDs) and LDs. Due to the internal electric field induced by discontinuities in spontaneous polarization of the wurtzite structure in the direction of the *c* axis, performance of LEDs and LDs can be reduced.[10] Thus, it is apparent that the elimination or lessening of polarization can be beneficial in these devices.[10] On the other hand, in the case of high-power or high-frequency applications, spontaneous polarization can be advantageous. Given this variation in outcome based on application, the ability to render wurtzite materials as polar, semipolar, or nonpolar (see Table 2.4) widens the capabilities and scope of these materials.

The aforementioned discontinuities in spontaneous polarization arise at the interface of mismatched materials. At this heterointerface, the variation in fixed polarization charges between the two materials leads to high two-dimensional electron gas (2DEG) concentration.[10] In high-power or high-frequency applications, typically AlGaN/GaN heterostructures, this effect is desirable.[10] However, for other applications such as LEDs and LDs, utilizing nonpolar and semipolar orientations that allow for adjustment of the degree of

TABLE 2.4

Polarization of Planes Commonly Used for Growth of Hexagonal Semiconductor Crystals

Polarization	Growth Plane
Polar	(0001), $(000\bar{1})$
Semipolar	$(10\bar{1}1)$[10], $(11\bar{2}2)$[10], $(10\bar{1}3)$[10]
Nonpolar	$(11\bar{2}0)$[10], $(10\bar{1}0)$[10]

polarization is desirable because in these cases, the induced electric field contributes to nonradiative recombination and reduced efficiency.[10]

2.2.8.4 Hexagonal Silicon Carbide

In 4H- and 6H-SiC, the {0001} planes are not equivalent. The (0001) face contains only Si atoms, whereas the (000$\bar{1}$) surface (sometimes called the (0001)C face) is made up entirely of C atoms. This leads to a number of observable differences in the chemical behavior of these faces. For example, rates of oxidation[11,12] and vapor phase epitaxial growth[13–15] are both faster on the (0001)C face than on the (0001)Si face.

2.3 Lattice Constants and Thermal Expansion Coefficients

In an unstrained cubic crystal, a single lattice constant, a, defines the length of the sides of the cubic unit cell. For a hexagonal crystal, there are two lattice constants, a and c. The former represents the distance from the sixfold rotation axis to a corner of the hexagonal base, and the latter represents the height of the unit cell. It is important to know the lattice constants of the substrate as well as the epitaxial layer, because they determine the lattice mismatch for heteroepitaxy. The lattice constants of elemental and binary semiconductors may be determined by x-ray diffraction experiments, with parts per million resolution (see Chapter 7 for a discussion of this).

Lattice constants increase with temperature due to normal thermal expansion. This can be an important effect in heteroepitaxy, which may take place at greatly elevated temperatures. This is especially true if the substrate and epitaxial layer have greatly different thermal expansion characteristics.

The linear thermal coefficient of expansion (TCE) α is defined as

$$\alpha = \frac{1}{a}\frac{\partial a}{\partial T}, \tag{2.51}$$

and has units of inverse Kelvins (K^{-1}). Typical values of the linear TCE are on the order of 10^{-6} to 10^{-5} K^{-1}, but the value itself depends on the temperature. Table 2.5 provides the lattice constants and linear TCEs for cubic semiconductor crystals.

The TCE is itself a function of temperature. Thus, the experimentally obtained thermal expansion characteristics are often fit to a polynomial:

$$\frac{\Delta a}{a} = A + BT + CT^2 + DT^3, \tag{2.52}$$

where the fractional change in lattice constant $\Delta a/a$ is with respect to 300K, and T is the absolute temperature in Kelvin. Thus, at a temperature T, the relaxed lattice constant for the crystal is given by

$$a(T) = a(300K)\left[1 + A + BT + CT^2 + DT^3\right]. \tag{2.53}$$

TABLE 2.5

Lattice Constants and Thermal Expansion Coefficients for Cubic Semiconductor Crystals

	a(300 K) (Å)	α(300 K) (10^{-6} K^{-1})	α(600 K) (10^{-6} K^{-1})	α(1000 K) (10^{-6} K^{-1})
C	3.56684[16]	1.0[17]	2.8[17]	4.4[17]
Si	5.43108[18]	2.6[17]	3.7[17]	4.4[17]
Ge	5.6576[19]	5.7[17]	6.7[17]	7.6[17]
α-Sn	6.4894[20]	4.7[20]	—	—
SiC (3C)	4.3596[21,22]	—	—	—
BN	3.615[23]	1.8[24]	3.7[24]	5.9[24]
BP	4.538[25]	—	—	—
Bas	4.777[25]	—	—	—
AlP	5.467[26]	—	—	—
AlAs	5.660[27]	—	—	—
AlSb	6.1357[28]	4.4[17]	—	—
GaP	5.4512[29]	4.7[24]	5.8[24]	—
GaAs	5.65325[30]	5.7[24]	6.7[24]	—
GaSb	6.0960[31]	6.1[17]	7.3[17]	—
InP	5.8690[28]	4.75[32]	—	—
InAs	6.0584[33]	5.19[34]	—	—
InSb	6.4794[35]	5.0[17]	6.1[17]	—
BeS	4.865[25]	—	—	—
BeSe	5.139[25]	—	—	—
BeTe	5.626[25]	—	—	—
ZnS	5.4105[36]	7.1[24]	8.6[24]	10.5[24]
ZnSe	5.6687[37]	7.1[24]	10.1[24]	—
ZnTe	6.1041[38]	8.8[24]	10.0[24]	—
CdTe	6.481[39]	5.0[17]	5.4[17]	—
β-HgS	5.851[36]	—	—	—
HgSe	6.084[36]	—	—	—
HgTe	6.461[40]	5.1[24]	—	—

From the polynomial characteristic the linear coefficient of thermal expansion may be determined as a function of temperature from

$$\alpha = \frac{1}{a}\frac{\partial a}{\partial T} = B + 2CT + 3DT^2. \tag{2.54}$$

The constants A, B, C, and D for cubic crystals are provided in Table 2.6. In the cases of GaAs and ZnS, two polynomials are given for two different temperature ranges.

For hexagonal crystals such as the III-nitrides, 4H- and 6H-SiC, and sapphire, the thermal expansion coefficients are different for the a and c lattice constants. Usually, both α_a, the thermal expansion coefficient for the lattice constant a along the $[11\bar{2}0]$ direction, and α_c, the thermal expansion coefficient for the lattice constant c along the $[0001]$ direction, are reported. Occasionally, α_m, the thermal expansion coefficient along the $[10\bar{1}0]$ direction, is also given.

Relatively little information has been published on the thermal expansion of the hexagonal crystals, and in some cases there are great disparities between the available data.

TABLE 2.6

Temperature Dependence of Thermal Expansion for Cubic Crystals

	A (10^{-2})	*B* ($10^{-6}\,K^{-1}$)	*C* ($10^{-9}\,K^{-2}$)	*D* ($10^{-12}\,K^{-3}$)	Temp. Range (K)
C[17]	–0.010	–0.591	3.32	–0.5544	(25–1650)
Si[17]	–0.071	1.887	1.934	–0.4544	(293–1600)
Ge[17]	–0.1533	4.636	2.169	–0.4562	(293–1200)
α-Sn[17]	–0.525	13.54	15.87	–2.896	(100–500)
BN[24]	–0.0013	–1.278	4.911	–0.8635	(293–1300)
AlSb[17]	–0.049	–2.997	22.43	–22.34	(40–350)
GaP[24]	–0.110	2.611	4.445	–2.023	(293–850)
GaAs[24]	–0.147	4.239	2.916	–0.936	(200–1000)
GaAs[41–45]	–0.155	4.411	2.70	–0.667	(293–643)
GaSb[17]	–0.138	3.051	66.02	–3.380	(100–800)
InSb[17]	–0.099	1.249	8.773	–5.260	(50–750)
ZnS[24]	–0.0863	–3.386	30.18	–29.21	(60–335)
ZnS[46–48]	–0.2173	6.628	2.15	–0.333	(293–820)
ZnSe[41,49–51]	–0.1566	3.044	8.25	–3.33	(293–870)
ZnTe[24]	–0.200	5.104	6.811	–3.104	(100–725)
CdTe[17]	–0.0980	1.624	7.176	–4.445	(100–700)
HgTe[24]	–0.504	9.772	42.66	–59.22	(50–300)

Note: $\Delta a/a = A + BT + CT^2 + DT^3$, where T is the absolute temperature.[a]

[a] These polynomial coefficients are absolute, so that $\Delta a/a$ is absolute. In the first edition of this book, the coefficients were given such that $\Delta a/a$ was in percent.

For example, the value of α_c(300 K) has been reported to be $3.2\times10^{-6}\,K^{-1}$ by Maruska and Tietjen,[52] $2.8\times10^{-6}\,K^{-1}$ by Leszczynski et al.,[53] and $5.8\times10^{-6}\,K^{-1}$ by Oshima et al.[54] This may be due, at least in part, to the different methods of preparation for the crystals examined. Also, the lack of experimental data for some materials reflects the difficulty in preparing bulk crystals for thermal expansion characterization. In light of these challenges, the values in Table 2.7 should be considered only as best estimates until more data become available.

It should be noted that the lattice constants depend on the native defect concentrations, including those of vacancies, interstitials, antisite defects, and their complexes, so that the lattice constants measured for high-purity semiconductors depend on the method of preparation. Moreover, the lattice constant may be modified by the presence of impurities, to an extent that is often directly proportional to the impurity concentration. Examples are O in GaN, which resides on the nitrogen sublattice and causes a decrease in the lattice parameters because it has a smaller tetrahedral radius than nitrogen;[66] Cl in ZnSe,[67] which causes a decrease in the lattice parameter; and Sb in Si,[68] which results in an increase in the lattice constant. Some workers have considered these modifications of the lattice constant to be associated with hydrostatic stress, especially in the case of GaN research,[69] although this is a description of convenience and not physically accurate.

The lattice constants of alloyed semiconductors such as binary alloys, ternaries, and quaternaries are often estimated by linear interpolation (Vegard's law[70]). Sometimes this is a matter of convenience, but in other cases lattice constants have not been measured and

TABLE 2.7

Lattice Constants and Thermal Expansion Coefficients for Hexagonal Crystals

	a (Å)	c (Å)	α_a (300 K) (10^{-6} K^{-1})	α_c (300 K) (10^{-6} K^{-1})	α_a (300 K) (10^{-6} K^{-1})	α_c (300 K) (10^{-6} K^{-1})
α-Al_2O_3	4.7592[55]	12.9916[55]	4.3	3.9	5.6	7.4
SiC (2H)	3.0763[21]	5.0480[21]	—	3.0[56]	—	3.4[56]
SiC (4H)	3.0730[38]	10.053	3.3[57]	3.2[57]	4.2[57]	3.8[57]
SiC (6H)	3.0806[21]	15.1173[21]	3.4[56]	3.3[56]	4.2[56]	3.9[56]
AlN	3.112[58]	4.9788[59]	4.4[59]	3.5[59]	—	—
BN	2.534[60]	4.191[60]				
GaN	3.1882(4)[61,62]	5.1855(5)[61,62]	3.1[53]	2.8[53]	4.7[53]	4.2[53]
InN	3.533[63]	5.693[63]	3.4[64]	2.7[64]	5.7[64]	3.7[64]
ZnS	3.8140	6.2576	8.4[65]	4.1[65]	9.9[65]	5.5[65]
ZnTe	4.27	6.99	—	—	—	—
CdS	4.1363[59]	6.7154[59]	4.3[59]	2.8[59]	—	—
CdSe	4.3007[59]	7.0113[59]	4.1[59]	2.8[59]	—	—
CdTe	4.57	7.47	—	—	—	—

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 1.562(3) is 1.562 ± 0.003.

reported over the entire compositional range of interest. For a binary alloy such as Ge_xSi_{1-x}, the lattice constant may be estimated by

$$a\left(Ge_xSi_{1-x}\right) = xa_{Ge} + \left(1-x\right)a_{Si}\ , \tag{2.55}$$

where a_{Ge} and a_{Si} are the relaxed lattice constants of Ge and Si, respectively. Ternary alloys can be formed by the mixing of II-VI or III-V semiconductors in which one element is common to both. For the ternary alloy $A_xB_{1-x}C$, which involves the random occupation of the cation sublattice by atoms A and B, the relaxed lattice constant may be estimated using Vegard's law as

$$a\left(A_xB_{1-x}C\right) = xa_{AC} + (1-x)\,a_{BC}\,, \tag{2.56}$$

where a_{AC} and a_{BC} are the relaxed lattice constants of the binary semiconductors AC and BC, respectively. For example, the relaxed lattice constant of the ternary $In_xGa_{1-x}As$ may be found by

$$a\left(In_xGa_{1-x}As\right) = xa_{InAs} + \left(1-x\right)a_{GaAs}\ , \tag{2.57}$$

where a_{InAs} and a_{GaAs} are the relaxed lattice constants of InAs and GaAs, respectively. Ternary alloys sometimes involve mixing on the anion sites, as well as in the ternary AB_yC_{1-y}. An example is $GaAs_yP_{1-y}$, utilized in red LEDs. Here the relaxed lattice constant is approximately

$$a\left(AB_yC_{1-y}\right) = ya_{AB} + \left(1-y\right)a_{AC}. \tag{2.58}$$

Occasionally, linear interpolation is not sufficiently accurate, and the second-order term must be included through the use of a "bowing parameter." Here, the lattice constant for a ternary with mixing on the cation sites is modified to

$$a\left(A_xB_{1-x}C\right) = xa_{AC} + \left(1-x\right)a_{BC} - \delta x(1-x), \tag{2.59}$$

where δ is the bowing parameter of the ternary alloy.

Quaternary alloys are also of technological importance, and these contain four types of atoms. Alloying may occur on the cation sublattice, the anion sublattice, or on both, leading to three types of quaternary materials. In type I, the alloying is all on the cation sublattice, so the quaternary $A_xB_yC_{1-x-y}D$ is a mixture of the semiconductors AD, BD, and CD. An example is $Al_xIn_yGa_{1-x-y}P$, used in red, orange, and yellow LEDs, and the lattice constant for this type is

$$a\left(A_xB_yC_{1-x-y}D\right) = xa_{AD} + ya_{BD} + \left(1-x-y\right)a_{CD}. \tag{2.60}$$

In a type II quaternary, alloying occurs on both the cation and anion sublattices, and thus the quaternary $A_xB_{1-x}C_yD_{1-y}$ is a mixture of the semiconductors AC, AD, BC, and BD. An example is $Zn_xMg_{1-x}S_ySe_{1-y}$, used in II-VI blue-green LEDs and LDs.[71] Here the lattice constant may be estimated from[72]

$$a\left(A_xB_{1-x}C_yD_{1-y}\right) = xya_{AC} + x\left(1-y\right)a_{AD} + \left(1-x\right)ya_{BC} + \left(1-x\right)\left(1-y\right)a_{BD}. \tag{2.61}$$

In the type III quaternary, the alloying is all on the anion sublattice, so the quaternary $AB_xC_yD_{1-x-y}$ is a mixture of the semiconductors AB, AC, and AD, with a common cation A. An example is $InAs_xSb_yP_{1-x-y}$, used in infrared LDs in the wavelength range of 2.7–3.6 μm.[73] For the quaternary with anion mixing the relaxed lattice constant may be estimated by

$$a\left(AB_xC_yD_{1-x-y}\right) = xa_{AB} + ya_{AC} + \left(1-x-y\right)a_{AD}. \tag{2.62}$$

As with the binary or ternary alloys, nonlinearities may be accounted for by retaining second-order terms through the use of bowing parameters. However, this is more common in the use of Vegard's law for the determination of energy gaps.

Occasionally Vegard's law for a quaternary is given in terms of the constants for the constituent ternaries, but this is more common in the case of bandgap estimation. For example, to determine the bandgap differences in the quaternary material systems, it may be beneficial to interpolate from the experimental bandgap energies of the ternary constituents[74] and the expression may be rewritten as

$$E^g\left(A_xB_{1-x}C_yD_{1-y}\right) = \frac{1}{x(1-x)+y(1-y)}\left[\begin{array}{l} x(1-x)\left[\begin{array}{l} yE^g_{A_xB_{1-x}C} \\ +(1-y)E^g_{A_xB_{1-x}D}\end{array}\right] \\ +y(1-y)\left[\begin{array}{l} xE^g_{A_yC_{1-y}D} \\ +(1-x)E^g_{B_yC_{1-y}D}\end{array}\right]\end{array}\right]. \tag{2.63}$$

2.4 Elastic Properties

Heteroepitaxial semiconductors typically contain elastic strains, due to lattice mismatch and thermal expansion mismatch. These strains affect the properties of semiconductor devices in diverse ways. For example, strain can change the band structure of a semiconductor, and the energy gap in particular. In Si metal oxide semiconductor field effect transistors (MOSFETs), tensile biaxial strain can improve the electron mobility,[75] while compressive biaxial strain in SiGe MOSFETs can increase the hole mobility.[76] Built-in strains can also promote the motion of dislocations during the operation of injection lasers, thus causing catastrophic device failure. This section introduces the basic theory of how semiconductor crystals respond to stresses. Special emphasis is given to tetragonal distortion and elastic strain energies in mismatched heteroepitaxial layers, because these concepts are used repeatedly in Chapters 5 and 6.

2.4.1 Infinitesimal Strain Theory

The development of strain theory is based on continuum mechanics and considers deformations in the natural structure of a crystal due to an applied stress (see Figure 2.12). The key assumption here is that the displacement of the natural structure is much smaller than the lattice parameters describing it, so that linear relationships may be applied.

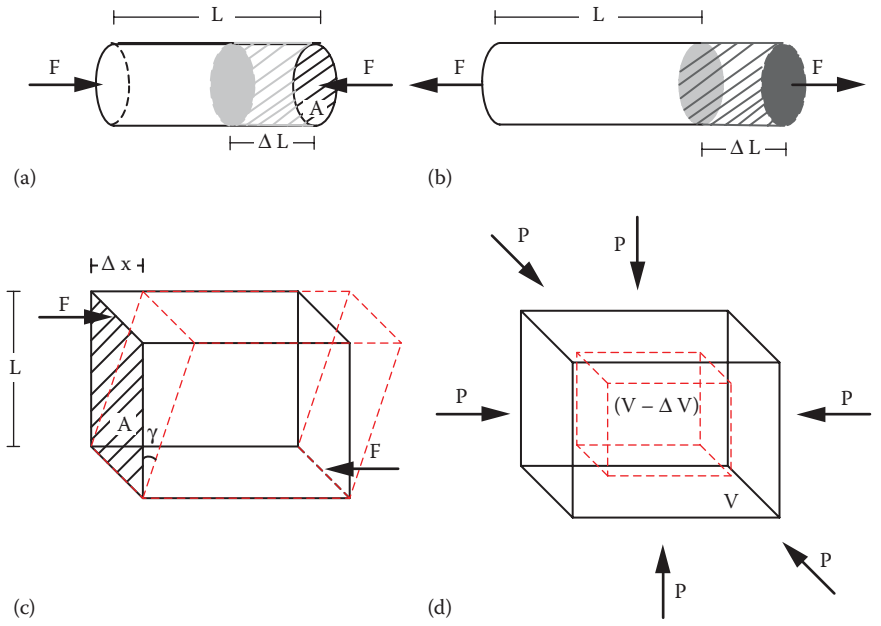


FIGURE 2.12
Stress–strain relations. (a) Compressive applied force to the normal side surfaces of a cylinder. (b) Change in cylinder length due to a tensile applied force. (c) Shear applied forces to a tetragonal crystal structure, along with its planar displacement. (d) Change in volume due to a hydrostatic pressure.

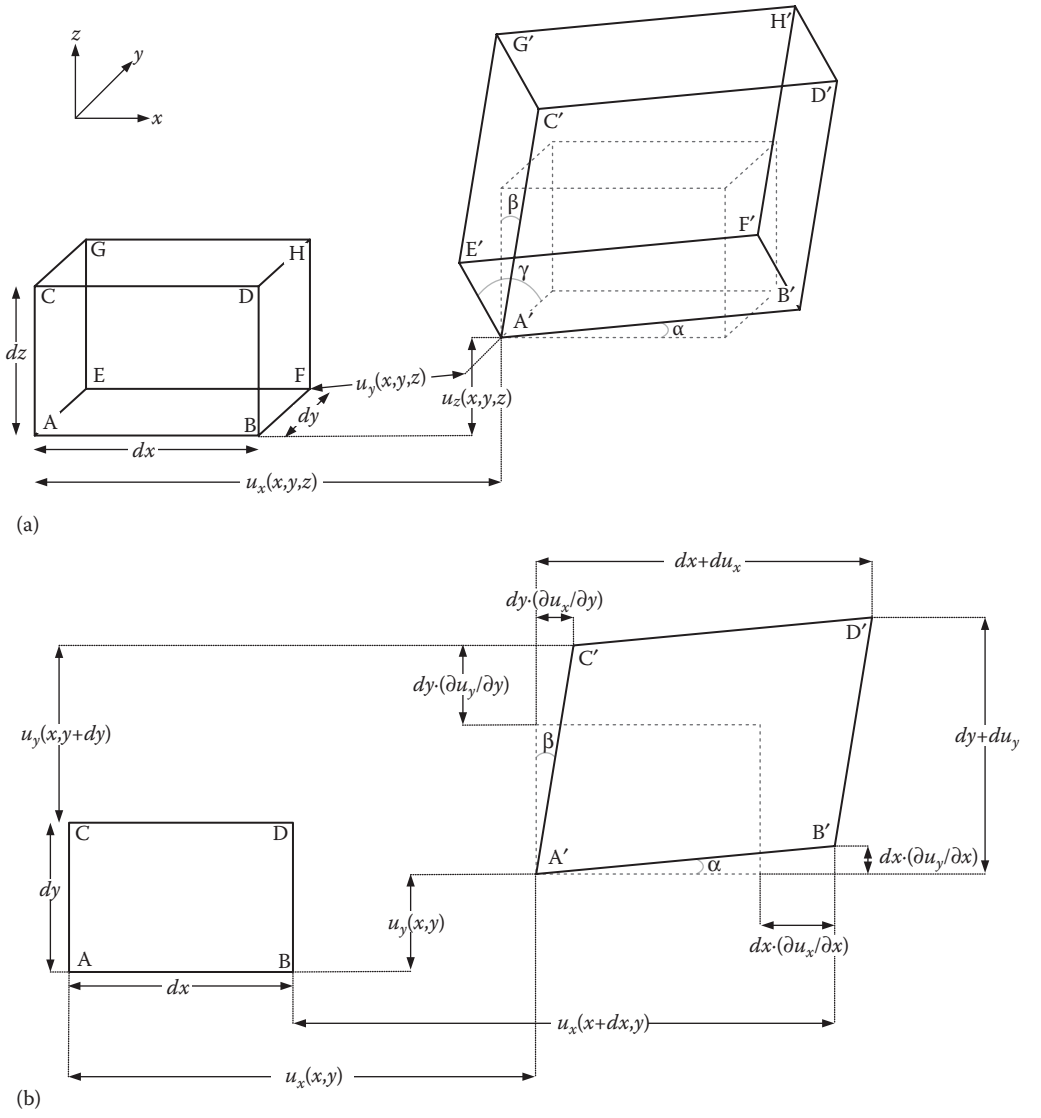


FIGURE 2.13
(a) Distortion of a 3-D solid body. (b) Representation of point displacement in two dimensions due to an applied stress.

Figure 2.13a shows a 3-D body in its natural state (described by its characteristic dimensions). Under an applied stress, we can define the distortions of the natural structure by giving the displacement of each point from its relaxed position. For clarity, we can confine the distortion of the body to two dimensions, as shown in Figure 2.13b. We can consider the rectangle described by points ABCD displaced to a distorted structure (rhombus) described by points A′B′C′D′. If we consider the displacement of the sides AB and AC to A′B′ and A′C′, we can write the lengths of both vectors $\overline{A'B'}$ and $\overline{A'C'}$ as

$$\begin{aligned}\overline{A'B'} &= \sqrt{\left(dx+\frac{\partial u_x}{\partial x}dx\right)^2+\left(\frac{\partial u_y}{\partial x}dx\right)^2} \\ &= dx\sqrt{1+2\frac{\partial u_x}{\partial x}+\left(\frac{\partial u_x}{\partial x}\right)^2+\left(\frac{\partial u_y}{\partial x}\right)^2},\end{aligned} \tag{2.64}$$

and

$$\begin{aligned}\overline{A'C'} &= \sqrt{\left(\frac{\partial u_x}{\partial y}dy\right)^2+\left(dy+\frac{\partial u_y}{\partial y}dy\right)^2} \\ &= dy\sqrt{1+2\frac{\partial u_y}{\partial y}+\left(\frac{\partial u_y}{\partial y}\right)^2+\left(\frac{\partial u_x}{\partial y}\right)^2}.\end{aligned} \tag{2.65}$$

Therefore, we can define the normal strain in the x direction by

$$\begin{aligned}\varepsilon_x &= \frac{\overline{A'B'}-\overline{AB}}{\overline{AB}} = \frac{\sqrt{dx^2\left(1+2\frac{\partial u_x}{\partial x}+\left(\frac{\partial u_x}{\partial x}\right)^2+\left(\frac{\partial u_y}{\partial x}\right)^2\right)}-dx}{dx} \\ &= \sqrt{\left(2\frac{\partial u_x}{\partial x}+\left(\frac{\partial u_x}{\partial x}\right)^2+\left(\frac{\partial u_y}{\partial x}\right)^2\right)}.\end{aligned} \tag{2.66}$$

In a similar fashion, the normal strain in the y direction is given by

$$\begin{aligned}\varepsilon_y &= \frac{\overline{A'C'}-\overline{AC}}{\overline{AC}} = \frac{\sqrt{dx^2\left(1+2\frac{\partial u_y}{\partial y}+\left(\frac{\partial u_y}{\partial y}\right)^2+\left(\frac{\partial u_x}{\partial y}\right)^2\right)}-dx}{dx} \\ &= \sqrt{\left(2\frac{\partial u_y}{\partial y}+\left(\frac{\partial u_y}{\partial y}\right)^2+\left(\frac{\partial u_x}{\partial y}\right)^2\right)}.\end{aligned} \tag{2.67}$$

For very small displacements, we can ignore the higher-order terms in Equations 2.64 and 2.65 and the length of the new vectors may be given as

$$\overline{A'B'} \approx dx\left(1+\frac{\partial u_x}{\partial x}\right) \tag{2.68}$$

and

$$\overline{A'C'} \approx dy\left(1+\frac{\partial u_y}{\partial y}\right). \tag{2.69}$$

The approximate first-order expressions for the normal strains reduce to

$$\varepsilon_x \approx \frac{\partial u_x}{\partial x}, \tag{2.70}$$

and

$$\varepsilon_y = \frac{\partial u_y}{\partial y}. \tag{2.71}$$

Looking at Figure 2.13, we can also define the displacements in the x and y directions as

$$du_x = \frac{\partial u_x}{\partial x}dx + \frac{\partial u_x}{\partial y}dy, \tag{2.72}$$

and

$$du_y = \frac{\partial u_y}{\partial x}dx + \frac{\partial u_y}{\partial y}dy. \tag{2.73}$$

By defining the following quantities,

$$e_{11} = \frac{\partial u_x}{\partial x}, e_{12} = \frac{\partial u_x}{\partial y}, e_{21} = \frac{\partial u_y}{\partial x}, e_{22} = \frac{\partial u_y}{\partial y}, \tag{2.74}$$

we can rewrite the displacement as

$$du_i = e_{ij}dx_j \quad (j = 1,2). \tag{2.75}$$

The relative displacement tensor e_{ij} is important because it gives information on the change in unit length and the rotation of the lattice vector. From Figure 2.13b, we can see that e_{11} gives the extension of the distorted basis vector per unit length resolved along the undistorted lattice vector of the crystal in the x direction, while e_{21} gives the counterclockwise rotation of the same vector. Similarly, e_{22} gives the extension of the distorted basis vector per unit length resolved along the undistorted lattice vector of the crystal in the y direction, while e_{12} gives the clockwise rotation of the same vector. However, the tensor e_{ij} does not provide an accurate measure of strain because it may be possible to have nonzero components in e_{ij} without there being any distortion on the crystal. This special case involves problems where only rotations are available. Therefore, to remove the component of rotation from the relative displacement tensor e_{ij}, it is expressed as the sum of a symmetrical tensor ε_{ij} and an antisymmetric tensor ω_{ij} such that

$$e_{ij} = \varepsilon_{ij} + \omega_{ij}, \tag{2.76}$$

where $\varepsilon_{ij} = 1/2(e_{ij} + e_{ji})$ is referred to as the pure strain and $\omega_{ij} = 1/2(e_{ij} - e_{ji})$ details the amount of rotation. The simple shear strain is the change in angle between original and distorted lattice vectors. For the x-y plane, the simple shear strain Γ_{xy} is defined as the sum of two radian angles shown in Figure 2.13b where

$$\Gamma_{xy} = \alpha + \beta, \tag{2.77}$$

where

$$\alpha = \frac{dx\dfrac{\partial u_y}{\partial x}}{dx + dx\dfrac{\partial u_x}{\partial x}} = \frac{\dfrac{\partial u_y}{\partial x}}{1 + \dfrac{\partial u_x}{\partial x}}, \tag{2.78}$$

and

$$\beta = \frac{dy\dfrac{\partial u_x}{\partial y}}{dy + dy\dfrac{\partial u_y}{\partial y}} = \frac{\dfrac{\partial u_x}{\partial y}}{1 + \dfrac{\partial u_y}{\partial y}}. \tag{2.79}$$

By considering small rotations where $\alpha, \beta \ll 1$ radian, then $\tan\alpha = \alpha$ and $\tan\beta = \beta$, recognizing that at small displacements $\alpha = \partial u_y / \partial x$ and $\beta = \partial u_x / \partial y$. Therefore,

$$\Gamma_{xy} = \frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y}. \tag{2.80}$$

By interchanging x and y along with u_x and u_y, it can be shown that

$$\Gamma_{xy} = \Gamma_{yx}. \tag{2.81}$$

Therefore, the simple shear strain components for the 3-D case could be written by considering the relation of Equation 2.80 or the symmetric part of the relative displacement tensor e_{ij}, given by

$$\varepsilon_{ij} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} = \begin{bmatrix} \varepsilon_{xx} & \Gamma_{xy}/2 & \Gamma_{xz}/2 \\ \Gamma_{yx}/2 & \varepsilon_{yy} & \Gamma_{yz}/2 \\ \Gamma_{zx}/2 & \Gamma_{zy}/2 & \varepsilon_{zz} \end{bmatrix}$$
$$= \begin{bmatrix} e_{11} & (e_{11}+e_{21})/2 & (e_{13}+e_{31})/2 \\ (e_{21}+e_{12})/2 & e_{22} & (e_{23}+e_{32})/2 \\ (e_{31}+e_{13})/2 & (e_{32}+e_{23})/2 & \varepsilon_{33} \end{bmatrix}. \tag{2.82}$$

2.4.2 Hooke's Law

Elastic strains in semiconductor crystals are in response to applied stresses. An arbitrary elastic strain may be specified by six quantities. If α, β, and γ are the angles between the ***a***, ***b***, and ***c*** axes in the unstrained crystal, then one possible set of such quantities is Δα, Δβ, Δγ, Δ***a***, Δ***b***, and Δ***c***. Because of the mathematical difficulties imposed by nonorthogonal axes, it is customary instead to use the six strains ε_{ij} defined as follows. The notation for strain ε_{ij} corresponds to the response of the strain in the *i direction* from a field applied in the *j direction*.

Three orthogonal axes ***f***, ***g***, and ***h*** of unit length are chosen within the unstrained crystal with their origins fixed at a particular lattice point, as shown in Figure 2.14. After a small

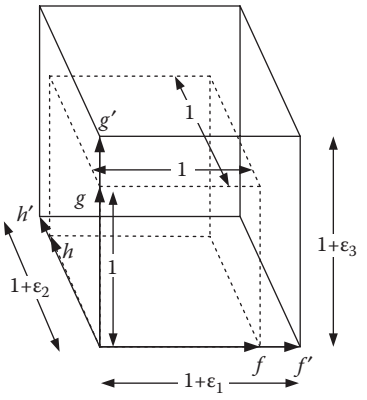


FIGURE 2.14
Volume dilation and orthogonal distortion in a 3-D crystal.

deformation of the crystal the axes are distorted in length and orientation to $\boldsymbol{f}'$, $\boldsymbol{g}'$, and $\boldsymbol{h}'$ such that

$$\boldsymbol{f}' = \left(1 + e_{xx}\right)\boldsymbol{f} + e_{xy}\boldsymbol{g} + e_{zx}\boldsymbol{h}, \tag{2.83}$$

$$\boldsymbol{g}' = e_{xy}\boldsymbol{f} + \left(1 + e_{yy}\right)\boldsymbol{g} + e_{yz}\boldsymbol{h}, \tag{2.84}$$

and

$$\boldsymbol{h}' = e_{zx}\boldsymbol{f} + e_{zy}\boldsymbol{g} + \left(1 + e_{zz}\right)\boldsymbol{h}. \tag{2.85}$$

The fractional changes in length of the $\boldsymbol{f}$, $\boldsymbol{g}$, and $\boldsymbol{h}$ axes are, to first order, given by

$$\varepsilon_{xx} \approx e_{xx}, \tag{2.86}$$

$$\varepsilon_{yy} \approx e_{yy}, \tag{2.87}$$

and

$$\varepsilon_{zz} \approx e_{zz}. \tag{2.88}$$

The shear strains,* or those strains related to the changes in α, β, and γ, are to first order

$$\varepsilon_{xy} = \boldsymbol{f}' \cdot \boldsymbol{g}' \approx e_{yx} + e_{xy}, \tag{2.89}$$

$$\varepsilon_{yz} = \boldsymbol{g}' \cdot \boldsymbol{h}' \approx e_{zy} + e_{yz}, \tag{2.90}$$

* In some references, the quantities ε_{xy}, ε_{yz}, and ε_{zx} are referred to as "engineering shear strains." They are approximately twice the simple shear strains e_{yx}, e_{xy}, e_{zy}, e_{yz}, e_{zx}, and e_{xz}. Thus, $\varepsilon_{xy} = e_{yx} + e_{xy} \approx 2e_{xy}$, $\varepsilon_{yz} = e_{zy} + e_{yz} \approx 2e_{yz}$, and $\varepsilon_{zx} = e_{zx} + e_{xz} \approx 2e_{zx}$. Engineering shear strains will be used throughout this book.

and

$$\varepsilon_{zx} = \boldsymbol{h}' \cdot \boldsymbol{f}' \approx e_{zx} + e_{xz} \,. \tag{2.91}$$

Stresses are deformational forces applied to the crystal, per unit area. Similarly, we will define the stress component σ_{ij} as a force applied in the i direction to a plane with its normal in the j direction.

2.4.2.1 Hooke's Law for Isotropic Materials

Hooke's law states that the strain components are linear combinations of the stress components. In an isotropic material, the physical properties are independent of direction. Therefore, Hooke's law takes on a simple form involving only two independent variables. In compliance form, Hooke's law for the isotropic medium is

$$\begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{bmatrix} = \frac{1}{E} \begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2+2\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 2+2\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 2+2\nu \end{bmatrix} \begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix}, \tag{2.92}$$

where E is the Young's modulus and ν is the Poisson ratio. These may be treated simply as material constants for our purposes here; they are described with more detail in Section 2.4.3.

The stresses may also be written as linear combinations of the strains. In stiffness form, Hooke's law for an isotropic medium is

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2}-\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}-\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}-\nu \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{bmatrix}, \tag{2.93}$$

2.4.2.2 Hooke's Law for Cubic Crystals

Cubic crystals are anisotropic in their elastic properties. Nonetheless, it is possible to greatly simplify Hooke's law by considerations of cubic symmetry. If the x, y, and z axes coincide with the [100], [010], and [001] directions in the cubic crystal, respectively, then Hooke's law in compliance form may be written as

$$\begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{12} & 0 & 0 & 0 \\ S_{12} & S_{11} & S_{12} & 0 & 0 & 0 \\ S_{12} & S_{12} & S_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & S_{44} \end{bmatrix} \begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} \tag{2.94}$$

or

$$\varepsilon = \boldsymbol{S}\sigma\,, \tag{2.95}$$

where the S_{ij} are the elastic compliance constants and $\boldsymbol{S}$ is the compliance matrix. Only three independent constants are needed as a consequence of the cubic symmetry.

In stiffness form, Hooke's law for a crystal with cubic symmetry is

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{bmatrix} \tag{2.96}$$

or

$$\sigma = \boldsymbol{C}\varepsilon\,, \tag{2.97}$$

where $\boldsymbol{C}$ is the stiffness matrix and the C_{ij} are the elastic stiffness constants, in units of force per area. Here, too, it is assumed that the x, y, and z axes coincide with the [100], [010], and [001] directions in the cubic crystal. The matrix Equation 2.96 applies in the general case. The Poisson ratio and the Young's modulus may also be used in heteroepitaxy as long as their dependence on the crystal direction is taken into account.

For cubic crystals, the compliance and stiffness constants are related by

$$C_{11} = \frac{S_{11} + S_{12}}{(S_{11} - S_{12})(S_{11} + 2S_{12})}, \tag{2.98}$$

$$C_{12} = \frac{-S_{12}}{(S_{11} - S_{12})(S_{11} + 2S_{12})}, \tag{2.99}$$

$$C_{44} = \frac{1}{S_{44}}\,, \tag{2.100}$$

$$S_{11} = \frac{C_{11} + C_{12}}{(C_{11} - C_{12})(C_{11} + 2C_{12})}, \tag{2.101}$$

$$S_{12} = \frac{-C_{12}}{(C_{11} - C_{12})(C_{11} + 2C_{12})}, \tag{2.102}$$

and

$$S_{44} = \frac{1}{C_{44}}. \tag{2.103}$$

Elastic constants of cubic crystals are often determined from acoustic measurements.[77] In these experiments ultrasonic pulses are generated in the crystal by a quartz transducer. The pulse traverses the crystal, is reflected by the back face, and returns. From the time elapsed, the velocity of propagation is determined. The measurement of three different wave modes allows calculation of all three unique elastic constants for a cubic crystal.

Table 2.8 provides the elastic stiffness constants C_{ij} for a number of cubic semiconductor crystals. Scarce elastic constant data are available in the literature for ternary and quaternary alloy layers. For a lack of a better approach, linear interpolation (Vegard's law) is often applied in these cases. However, there have been theoretical predictions of significant departures from linearity in some cases, including $In_{1-x}Ga_xSb$,[78] $Cd_{1-x}Zn_xTe$,[79] and $Si_{1-x-y}Ge_xC_y$.[80] Experimental data also suggest significant departures from linearity in the dilute nitride semiconductor $GaAs_{1-y}N_y$.[81]

TABLE 2.8

Elastic Stiffness Constants of Cubic Semiconductor Crystals at Room Temperature, in units of GPa (1 GPa = 10^{10} dyn/cm^2)

	C_{11}	C_{12}	C_{44}
C[82]	107.6	12.52(23)	57.74(14)
Si[83]	160.1	57.8	80.0
Ge[84]	124.0	41.3	68.3
α-Sn[85]	69.0	29.3	36.2
SiC (3C)[86]	352	120	232.9
AlN (ZB)[87]	322	156	138
AlP[88]	132	63.0	61.5
AlAs[88]	125	53.4	54.2
AlSb[89,90]	87.69(20)	43.41(20)	40.76(8)
GaN (ZB)[87]	325	142	147
GaP[91]	140.50(28)	62.03(24)	70.33(7)
GaAs[92]	118.4(3)	53.7(16)	59.1(2)
GaSb[93]	88.50	40.40	43.30
InP[94]	102.2	57.6	46.0
InAs[95]	83.29	45.26	39.59
InSb[96]	65.92(5)	35.63(6)	29.96(3)
ZnS[97]	104.62(5)	65.33(6)	46.50(12)
ZnSe[98]	87.2(8)	52.4(8)	39.2(4)
ZnTe[99]	71.3	40.7	31.2
CdTe[100]	53.3	36.5	20.44
β-HgS[101]	81.3	62.2	26.4
HgSe[101]	69.0	51.9	23.3
HgTe[102]	53.61	36.60	21.23

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 12.52(23) represents 12.52 ± 0.023.

2.4.2.3 Hooke's Law for Hexagonal Crystals

For a crystal with hexagonal symmetry (wurtzite semiconductor or hexagonal SiC), there are six distinct elastic stiffness constants, of which five are independent. Assuming that the z axis is aligned with the c axis of the hexagonal unit cell, Hooke's law can be written in stiffness form as

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{zx} \\ \varepsilon_{xy} \end{bmatrix}. \tag{2.104}$$

Elastic stiffness constants for hexagonal crystals may be determined from acoustic measurements as in the case of cubic crystals. In some cases, the resonance method[103,104] is used to determine the piezoelectric and elastic stiffness constants for piezoelectric hexagonal crystals such as 4H-SiC and 6H-SiC. In these experiments, only a subset of the elastic stiffness constants may be determined, depending on the orientation of the piezoelectric transducer, which is cut from a single crystal of the material under test.

Tables 2.9 through 2.12 provide the published elastic stiffness constants C_{ij} for some hexagonal semiconductors. It should be noted that only the first five of these constants are independent. C_{66} is not always reported but may be calculated from[105]

$$C_{66} = \frac{C_{11} - C_{12}}{2}. \tag{2.105}$$

2.4.3 Elastic Moduli

Some elastic properties that are useful in heteroepitaxy are the Young's modulus E, the biaxial modulus Y, the shear modulus G, the Poisson ratio ν, and the biaxial relaxation

TABLE 2.9

Elastic Stiffness Constants of 4H- and 6H-SiC at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

	4H-SiC	6H-SiC
Elastic Constants	**Kamitani et al.[106]**	**Kamitani et al.[106]**
C_{11}	507(4)	501(4)
C_{12}	111(5)	111(5)
C_{13}	—	52(9)
C_{33}	547(7)	553(4)
C_{44}	159(4)	163(4)
C_{66}	198	195

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 390(15) represents 390 ± 15.

TABLE 2.10

Elastic Stiffness Constants of Wurtzite GaN at Room Temperature, in Units of GPa ($1\ \text{GPa} = 10^{10}\ \text{dyn/cm}^2$)

Elastic Constants	Recommended Values	Polian et al.[107]	Deger et al.[108]	Deguchi et al.[109]	Davydov et al.[110]	Savastenko and Sheleg[111]
C_{11}	353	390(15)	370	373	315	296
C_{12}	135	145(20)	145	141	118	120
C_{13}	104	106(20)	110	80.4	96	158
C_{33}	367	398(20)	390	387	324	267
C_{44}	91	105(10)	90	93.6	88	24
C_{66}	110	123(10)	112	118	99	88

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 507(4) represents 507 ± 4.

TABLE 2.11

Elastic Stiffness Constants of Wurtzite AlN at Room Temperature, in Units of GPa ($1\ \text{GPa} = 10^{10}\ \text{dyn/cm}^2$)

Elastic Constants	Recommended Values	Deger et al.[108]	Davydov et al.[110]	McNeil et al.[112]	Tsubouchi et al.[113]	Davydov and Solomonov et al.[87]
C_{11}	397	410	419	411	345	369
C_{12}	145	140	177	149	125	145
C_{13}	113	100	140	99	120	120
C_{33}	392	390	392	389	395	395
C_{44}	118	120	110	125	125	96
C_{66}	128	135	121	131	131	112

TABLE 2.12

Elastic Stiffness Constants of Wurtzite InN at Room Temperature, in Units of GPa ($1\ \text{GPa} = 10^{10}\ \text{dyn/cm}^2$)

Elastic Constants	Recommended Values	Sheleg and Savastenko[63]	Kim et al.[114]	Wright[115]	Marmalyuk et al.[116]	Chisholm et al.[117]
C_{11}	250	190	271	223	257	297.5
C_{12}	109	104	124	115	92	107.4
C_{13}	98	121	94	92	70	108.7
C_{33}	225	182	200	224	278	250.5
C_{44}	54	9.9	46	48	68	89.4
C_{66}	70	43	74	54	82	95

constant R_B. The Young's modulus (also called the modulus of elasticity or the elastic modulus) is a measure of the stiffness of a material. It is defined as the ratio of stress to strain:

$$\text{Young's modulus} = E = \left.\frac{\text{stress}}{\text{strain}}\right|_{\text{uniaxial stress}}. \tag{2.106}$$

Usually, this definition for the Young's modulus is used with the assumption of a stress in one direction (uniaxial stress). For the case of biaxial stress commonly encountered in mismatched heteroepitaxy, we use the biaxial modulus, which is the ratio of the stress to strain for the biaxial case:

$$\text{Biaxial modulus} = Y = \left.\frac{\text{stress}}{\text{strain}}\right|_{\text{biaxial stress}}. \tag{2.107}$$

It should be noted, however, that the biaxial modulus is sometimes referred to as the Young's modulus in the literature. The shear modulus (also known as the "rigidity modulus") is defined as the ratio of the shear stress to shear strain:

$$\text{Shear modulus} = G = \frac{\text{shear stress}}{\text{shear strain}} \tag{2.108}$$

The Poisson ratio is defined as the ratio of the transverse contraction to the longitudinal extension, for a uniaxial tensile stress in the longitudinal direction. Thus,

$$\text{Poisson ratio} = \nu = -\left.\frac{\text{transverse strain}}{\text{longitudinal strain}}\right|_{\text{uniaxial stress}} \tag{2.109}$$

Typical semiconductor crystals have a Poisson ratio of 1/3. The Poisson ratio is nearly always positive, because the unit cell volume is approximately conserved in the strained crystal. The biaxial relaxation constant is analogous to the Poisson ratio, for the case of biaxial stress, so that

$$\text{Biaxial relaxation constant} = R_B = -\left.\frac{\text{in-plane strain}}{\text{out-of-plane strain}}\right|_{\text{biaxial stress}} \tag{2.110}$$

In the following sections, these elastic moduli are related to the elastic stiffness constants, and their values are tabulated for the crystals commonly used in heteroepitaxy.

2.4.3.1 Elastic Moduli for Cubic Crystals

For diamond and zinc blende crystals, the shear modulus is

$$G = \frac{(C_{11} - C_{12})}{2}. \tag{2.111}$$

If the growth plane is (001), the Young's modulus is

$$E(001) = \frac{(C_{11} + 2C_{12})(C_{11} - C_{12})}{(C_{11} + C_{12})} \tag{2.112}$$

and the Poisson ratio is

$$\nu(001) = \frac{C_{12}}{C_{11} + C_{12}}. \tag{2.113}$$

The biaxial modulus is given by

$$Y(001) = C_{11} + C_{12} - \frac{2C_{12}^2}{C_{11}} = \frac{E(001)}{1 - \nu(001)}, \tag{2.114}$$

and the biaxial relaxation constant is

$$R_B(001) = \frac{2C_{12}}{C_{11}}. \tag{2.115}$$

Elastic moduli for cubic semiconductors are given in Table 2.13.

2.4.3.2 Elastic Moduli for Hexagonal Crystals[118]

For wurtzite crystals or hexagonal SiC, the shear modulus is anisotropic, but may be derived from the tensor form of Hooke's law. If the growth plane is assumed to be (0001), then the Young's modulus is

$$E(0001) = C_{33} - \frac{2C_{13}^2}{(C_{11} + C_{12})}. \tag{2.116}$$

TABLE 2.13

Elastic Moduli of Cubic Semiconductor Crystals at 300 K

	G(GPa)	*E*(001) *(GPa)*	ν(001)	*Y*(001) *(GPa)*	R^B(001)
C	47	105	0.104	117	0.23
Si	51	129	0.265	176	0.72
Ge	41	103	0.25	138	0.67
α-Sn	19.8	52	0.30	73	0.85
SiC (3C)	116	290	0.25	390	0.68
AlN (ZB)	83	220	0.33	330	0.97
AlP	34	91	0.32	135	0.95
AlAs	36	93	0.30	133	0.85
AlSb	22	59	0.33	88	0.99
GaN (ZB)	92	240	0.30	340	0.87
GaP	39	102	0.31	148	0.88
GaAs	32	85	0.31	124	0.91
GaSb	24	63	0.31	92	0.91
InP	22	61	0.36	95	1.13
InAs	19.0	51	0.35	79	1.09
InSb	15.1	41	0.35	63	1.08
ZnS	19.6	54	0.38	88	1.25
ZnSe	17.4	48	0.38	77	1.20
ZnTe	15.3	42	0.36	66	1.14
CdTe	8.4	24	0.41	40	1.37
β-HgS	9.6	27	0.43	48	1.53
HgSe	8.6	24	0.43	43	1.50
HgTe	8.5	24	0.41	40	1.37

Note: These values were calculated using the elastic stiffness constants of Table 2.7; see that table for references.

TABLE 2.14

Elastic Moduli of Hexagonal Semiconductor Crystals at 300 K

	$E(0001)$ (GPA)	$\nu(0001)$	$Y(0001)$ (GPA)	$R^B(0001)$
6H-SiC	540	0.085	602	0.19
GaN[107]	320	0.21	430	0.57
AlN	340	0.21	480	0.58
InN	171	0.27	270	0.87

The Poisson ratio is

$$\nu(0001) = \frac{C_{13}}{C_{11} + C_{12}}, \tag{2.117}$$

and the shear modulus is given by[119]

$$G = \frac{E}{2(1+\nu)} = \frac{C_{33}(C_{11} + C_{12}) - 2}{2C_{13}}. \tag{2.118}$$

The biaxial modulus is

$$Y(0001) = C_{11} + C_{12} - \frac{2C_{13}^2}{C_{33}} = \frac{E}{1-\nu}, \tag{2.119}$$

and the biaxial relaxation constant is given by[120]

$$R_B(0001) = \frac{2C_{13}}{C_{33}}. \tag{2.120}$$

Values of elastic moduli for III-nitrides and 6H-SIC are given in Table 2.14.

2.4.4 Biaxial Stresses and Tetragonal Distortion in Cubic Crystals

For heteroepitaxial growth, we usually assume the case of biaxial stress. Using a Cartesian coordinate system, if growth proceeds along the z direction and the growth plane is the x-y plane, then the in-plane stresses applied by the substrate are equal:

$$\sigma_{xx} = \sigma_{yy} = \sigma_{||}. \tag{2.121}$$

Also, the out-of-plane stress is assumed to be zero:

$$\sigma_{zz} = \sigma_{\perp} = 0. \tag{2.122}$$

(The substrate does not constrain the epitaxial layer in the growth direction.) The shear stresses are assumed to be zero for growth on a low-index plane such as the (001) on a cubic

crystal or the (0001) on a hexagonal crystal. Also, it is usually assumed that the substrate is unstrained, because under most circumstances the substrate will be many times thicker (factor of ~100) than the epitaxial layer.

The stress tensor in the epitaxial layer is therefore given by

$$\sum = \begin{bmatrix} \sigma_{||} & 0 & 0 \\ 0 & \sigma_{||} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.123}$$

In the case of a biaxial stress applied to an (001) cubic crystal, the unit cell of the epitaxial layer becomes tetragonal with an in-plane lattice constant a and an out-of-plane lattice constant c. In this situation, referred to as *tetragonal distortion*,

$$\varepsilon_{||} = \varepsilon_{xx} = \varepsilon_{yy} = \frac{a - a_0}{a_0} \tag{2.124}$$

and

$$\varepsilon_{\perp} = \frac{c - a_0}{a_0}, \tag{2.125}$$

where a_0 is the relaxed (unstrained) lattice constant for the epitaxial layer. The in-plane strain is related to the biaxial stress by

$$\sigma_{||} = Y\varepsilon_{||} = \left(C_{11} + C_{12} - \frac{2C_{13}^2}{C_{33}} \right) \varepsilon_{||}, \tag{2.126}$$

where the constant of proportionality Y is the biaxial modulus described in the previous section. The in-plane and out-of-plane strains are related by

$$\varepsilon_{\perp} = -R_B \varepsilon_{||} = -\frac{2C_{12}}{C_{11}} \varepsilon_{||}, \tag{2.127}$$

where R_B is the biaxial relaxation constant. The strain tensor is therefore

$$\mathrm{E} = \begin{bmatrix} \sigma_{||}/Y & 0 & 0 \\ 0 & \sigma_{||}/Y & 0 \\ 0 & 0 & -\sigma_{||} R_B / Y \end{bmatrix}. \tag{2.128}$$

The equations above may be applied to pseudomorphic or partially relaxed heteroepitaxial layers, regardless of the presence or absence of thermal strain. Additionally, while there is some commonality in the equations between cubic and hexagonal crystals, the relationships for the biaxial modulus and biaxial relaxation constant are different. The following section details these differences.

2.4.5 Biaxial Stresses in Hexagonal Crystals

For heteroepitaxial growth of hexagonal materials on the (0001) *c*-plane, biaxial stresses applied in the plane of the interface result in a strain tensor with the following components:[118]

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a - a_0}{a_0} \tag{2.129}$$

and

$$\varepsilon_{zz} = \frac{c - c_0}{c_0} . \tag{2.130}$$

The relationship between the stress and strain can be described by[121]

$$\sigma_{xx} = \sigma_{yy} = \left(C_{11} + C_{12}\right)\varepsilon_{xx} + C_{13}\varepsilon_{zz} - e_{31}E_z \tag{2.131}$$

and

$$\sigma_{zz} = 2C_{13}\varepsilon_{xx} + C_{33}\varepsilon_{zz} - e_{33}E_z , \tag{2.132}$$

where E_z represents the electric field component for the polar axis (*c* axis) and e_{33} represents the piezoelectric contribution.[121] In some reported work, the piezoelectric and polarization effects have been neglected.

If the piezoelectric and polarization contributions are neglected, the stress–strain relationship reduces to

$$\sigma_{xx} = Y\varepsilon_{xx} = \left[C_{11} + C_{12} - \frac{2C_{13}^2}{C_{33}}\right]\varepsilon_{xx} \tag{2.133}$$

and

$$\sigma_{zz} = 0 . \tag{2.134}$$

The in-plane and out-of-plane strains are related by the biaxial relaxation constant,[122]

$$\varepsilon_{zz} = -R_B\varepsilon_{xx} = -\frac{2C_{13}}{C_{33}}\varepsilon_{xx} . \tag{2.135}$$

2.4.6 Strain Energy in Cubic Crystals

A load that produces a stress, acting on a crystal to deform it, does an amount of work per unit volume

$$\delta U = \sigma_{xx}\delta\varepsilon_{xx} + \sigma_{yy}\delta\varepsilon_{yy} + \sigma_{ZZ}\delta\varepsilon_{ZZ} + \sigma_{xy}\delta\varepsilon_{xy} + \sigma_{yz}\delta\varepsilon_{yz} + \sigma_{zx}\delta\varepsilon_{zx} . \tag{2.136}$$

Integrating the above expression, we can find the total strain energy per unit volume, which for the case of a cubic crystal is

$$U = \frac{C_{11}}{2}\left(\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2\right) + C_{12}\left(\varepsilon_{xx}\varepsilon_{yy} + \varepsilon_{yy}\varepsilon_{zz} + \varepsilon_{zz}\varepsilon_{xx}\right)$$
$$+ \frac{C_{44}}{2}\left(\varepsilon_{xy}^2 + \varepsilon_{yz}^2 + \varepsilon_{zx}^2\right) \quad . \tag{2.137}$$

For a biaxially strained heteroepitaxial layer, in which the shear terms vanish,

$$U = \frac{C_{11}}{2}\left(2\varepsilon_{||}^2 + \varepsilon_{\perp}^2\right) + C_{12}\left(\varepsilon_{||}^2 + 2\varepsilon_{||}\varepsilon_{\perp}\right)$$
$$= \varepsilon_{||}^2\left[C_{11} + C_{12} - \frac{2C_{12}^2}{C_{11}}\right]$$
$$= \varepsilon_{||}^2 Y \quad . \tag{2.138}$$

Therefore, the strain energy per unit area is

$$E_\varepsilon = \varepsilon_{||}^2 Y h \ , \tag{2.139}$$

where $\varepsilon_{||}$ is the in-plane strain, Y is the biaxial modulus, and h is the layer thickness. Early calculations of the critical layer thickness for heteroepitaxy involved balancing this strain energy with the energy of a grid of strain-relieving misfit dislocations.

2.4.7 Strain Energy in Nitride Semiconductors

The strain energy per unit area in a biaxially strained nitride semiconductor grown on the (0001) plane may be found in similar fashion and is given by

$$E_\varepsilon = \varepsilon_{||}^2 Y h = \varepsilon_{||}^2\left(C_{11} + C_{12} - \frac{2C_{13}^2}{C_{33}}\right)h$$
$$= \left[\left(C_{11} + C_{12}\right)\varepsilon_{xx}^2 + 2C_{13}\varepsilon_{xx}\varepsilon_{zz} + \frac{1}{2}C_{33}\varepsilon_{zz}^2\right]h \, . \tag{2.140}$$

2.5 Surface Free Energy

The growth mode of a nucleating heteroepitaxial layer is determined in large part by the properties of the surfaces and interfaces involved. The most important physical property is the surface free energy, defined as the reversible work done to create new surface area.

The surface free energies have been determined experimentally for only a few semiconductor crystals. In cases for which such data are available, the experimental errors are often quite large. These follow from the difficulties involved in surface free energy determinations. Most such efforts have involved the use of a fracture technique with natural cleavage planes. In these experiments, a precursor crack is introduced either by a steel wedge

(double-cantilever beam method[123]) or by an explosive electrical spark discharge (electrical spark discharge method[124]). The crack is expanded by application of a tensile force, and the relevant surface energy is determined using the Griffith criterion for crack propagation.[125] More recently, surface free energies have been determined for some semiconductors using the observed "equilibrium" shapes of facetted crystals (Bonzel method[126]). However, the accuracy of the values obtained by the Bonzel method depends on the assumed temperature dependence of the step free energy. Another approach, which is model independent, has been developed by Metois and Muller.[127] This involves observation of the equilibrium shapes for 3-D crystals and 2-D islands, as well as the statistical analysis of the thermal fluctuation for an isolated step. Data from these three experiments can be combined to find the surface free energies for the faces of the 3-D crystal.

Despite the progress with experimental methods, surface free energies are most often estimated by theoretical calculations. The most common approach involves use of the bond-breaking model applicable to covalent crystals.[128] In this model, the surface free energy is assumed to equal the bond strength B times the areal density of broken bonds for the crystal face. Then, in diamond and zinc blende crystals,

$$\gamma(001) = \frac{2B}{a^2},$$

$$\gamma(011) = \frac{2B}{a^2\sqrt{2}}, \text{ and}$$

$$\gamma(111) = \frac{2B}{a^2\sqrt{3}}, \tag{2.141}$$

where $\gamma(hkl)$ is the surface free energy of the (hkl) face, B is the bond energy, and a is the lattice constant. Surface free energies for the III-V semiconductors have been calculated using this model, using the bond strength B determined from the molar-atomic heat of sublimation, ΔH_s:

$$B = \frac{\Delta H_S}{2N_a}, \tag{2.142}$$

where N_a is the Avogadro constant, $6.022 \times 10^{23}\,\text{mol}^{-1}$.

Surface energies of the low-index faces may also be estimated based on knowledge of the lattice constant and elastic stiffness constants, which are more readily available in the literature.[125] Consider the strength of a cubic semiconductor rod having its long axis aligned with the [001] crystallographic direction. As the rod is stressed, elastic energy is stored until the rod breaks, at which time the elastic energy is converted to surface energy by the creation of two new (001) surfaces. The strength of the rod should not depend on its length, so we will suppose that all of the elastic energy is stored between two atomic planes.

The restraining force per unit area σ is supposed to have the form shown in Figure 2.15, which may be approximated by a sinusoid,

$$\sigma = K \sin\left[\frac{\pi(x - x_0)}{W}\right], \tag{2.143}$$

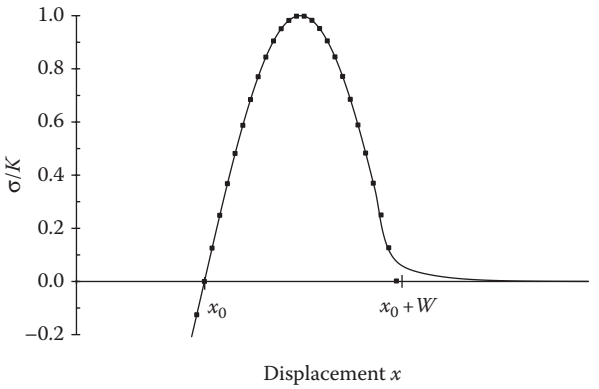


FIGURE 2.15
Restraining force per unit area σ vs. displacement x, approximated by a sinusoid.

where σ is the stress, x is the displacement, and K, x_0, and W are constants. x_0 is the relaxed (001) plane spacing, at which the restraining force is zero.

For a zinc blende crystal, $x_0 = a/4$. W is a measure of the range of interatomic forces. A rough estimate for W is between one and two times the interplanar spacing; the geometric mean $a/\sqrt{8}$ will be used somewhat arbitrarily. The constant K may be found as follows. If E is the Young's modulus for the [001] direction, then

$$E = x_0 \frac{d\sigma}{dx} . \tag{2.144}$$

Differentiating for small displacements ($x \approx x_0$), we find

$$K = \frac{E\sqrt{2}}{\pi} , \tag{2.145}$$

and the restraining force per unit area is

$$\sigma = \left(\frac{E\sqrt{2}}{\pi} \right) \sin\left[\frac{(x - x_0)\pi\sqrt{8}}{a} \right] . \tag{2.146}$$

Using this result, the (001) surface energy may be calculated as follows. The work done to break the rod is considered equal to the surface energy of the two surfaces created. Then,

$$\int_{x_0}^{x_0+W} \sigma dx = 2\gamma(001) \tag{2.147}$$

or

$$\gamma(001) = \frac{Ea}{2\pi^2} . \tag{2.148}$$

In terms of the elastic stiffness constants, the surface energy for the (001) plane of a cubic semiconductor crystal is given by

$$\gamma(001) = \frac{a(C_{11} - C_{12})(C_{11} + 2C_{12})}{2\pi^2 (C_{11} + C_{12})} . \tag{2.149}$$

The surface energies of the other low-index faces may be estimated using this relationship and the bond-breaking model. Thus, for a cubic semiconductor crystal,

$$\gamma(011) = \frac{\gamma(001)}{\sqrt{2}} \tag{2.150}$$

and

$$\gamma(111) = \frac{\gamma(001)}{\sqrt{3}} . \tag{2.151}$$

Table 2.15 summarizes values of the surface energies for low-index faces of cubic semiconductors. The values in parentheses were determined experimentally, and the values in square brackets were calculated using the heat of sublimation (Equation 2.142). All

TABLE 2.15

Surface Free Energies for the Low-Index Faces of Cubic Crystals at 300 K, in erg/cm^2

	$\gamma(001)$	$\gamma(011)$	$\gamma(111)$
C	1900	1300	1100
Si[129]	3600	2500 (1900)	2100 (1140)
Ge	3000	2100	1200
α-Sn	1700	1200	1000
AlP[130]	2500 [3400]	1800 [2400]	1400 (2000)
AlAs[130]	2700 [2600]	1900 [1800]	1600 [1500]
AlSb[130]	1800 [1900]	1300 [1300]	1000 [1100]
GaP[129]	2800	2000 (1900)	1600
GaAs[129]	2400	1700 (860)	1400
GaSb[130]	2000 [1600]	1400 [1100]	1200 [910]
InP[130]	1800 [1900]	1300 [1300]	1000 [1100]
InAs[130]	1600 [1400]	1100 [1000]	900 [840]
InSb[130]	1300 [1100]	900 [750]	800 [600]
ZnS	1500	1100	900
ZnSe	1400	1000	800
ZnTe	1300	900	800
CdTe	450	320	260
HgS	800	600	500
HgSe	800	600	500
HgTe	800	600	500

Note: The values in parentheses were determined experimentally, and the values in square brackets were calculated using the heat of sublimation. Other values were estimated using the elastic moduli.

TABLE 2.16

Surface Free Energies for Faces of Wurtzite Crystals at 300K, in erg/cm^2

Surface	γ
(0001)-AlN	5850[131]
$(10\bar{1}0)$-GaN (Ga)	1760[132]
(0001)-ZnO (Zn)	2490[133]
$(000\bar{1})$ -ZnO (O)	1350[134]
(0001)-CdSe (Se)	2400
$(000\bar{1})$-CdSe (Cd)	2080
(0001)-CdSe (Cd)	1440
$(000\bar{1})$ -CdSe (Se)	960

Note: All referenced values were determined using DFT. Face termination is indicated in parentheses following the material.

other values were calculated using the elastic stiffness constants and the lattice constant (Equations 2.149 through 2.151).

Both calculated and experimental surface energies for wurtzite materials remain scarce. The polarized nature of the wurtzite crystal complicates surface energy determinations. The primary method of calculation for the surface energies of wurtzite materials is rooted in density functional theory (DFT). Table 2.16 provides a compilation of surface energies calculated for a few wurtzite materials.

2.6 Dislocations

For partially relaxed layers that are greater than the critical layer thickness, misfit dislocations are produced at the interface to relieve some of the mismatch strain. Associated with these misfit dislocations are threading dislocations, which run through the thickness of the heteroepitaxial layer. An understanding of dislocations and their origin is important for the application of heteroepitaxy, because these defects tend to degrade the performance of devices.

Dislocations are linear defects, along which the interatomic bonding is disturbed relative to the case of a perfect crystal. In the *core* of the dislocation, along its line, there are dangling bonds and large local strains that exceed the limits of continuum elasticity theory. Surrounding the core is a strained region, in which the interatomic bonds are distorted by small amounts and Hooke's law can be considered to hold.

Some of the basic features of dislocations may be illustrated using 2-D bubble rafts. Figure 2.16 is a photograph of such a bubble raft, which contains a single dislocation. The dislocation is a point defect in the 2-D lattice and results in an extra half line of bubbles in the lower part of the raft. Matthews[134] has used lattice-mismatched bubble rafts to create models of misfit dislocations.

In a 3-D crystal lattice, the dislocation is a linear defect that may be associated with an extra half plane of atoms. Individual dislocations may be imaged using high-resolution transmission electron microscopy (HRTEM) and fast Fourier transforms. Figure 2.17 shows a filtered HRTEM image of a dislocation in heteroepitaxial AlN/6H-SiC (0001). Here, the misfit

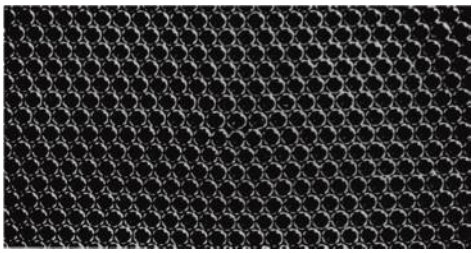


FIGURE 2.16
Dislocation in a bubble raft. The dislocation near the center of the photograph is associated with an extra half line of bubbles in the lower portion of the image. It is most easily seen by viewing the page with a shallow angle. (Photo courtesy of the University of Cambridge, Cambridge, UK, from the DoITPoMS web site. www.msn.cam.ac.uk. With permission.)

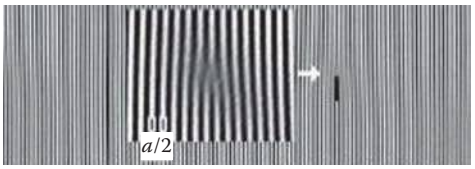


FIGURE 2.17
Filtered HRTEM image of a misfit dislocation in heteroepitaxial Al/6H-SiC (0001). Associated with the dislocation is an extra half plane of atoms in the 6H-SiC substrate. The $11\bar{2}0$-filtered image was obtained by a fast Fourier transform of the HRTEM cross-sectional image. (Reprinted from X. R. Huang et al., *Phys. Rev. Lett*., 95, 86101 [2005]. With permission. Copyright 2005, American Physical Society.)

dislocation manifests as a linear defect in the plane of the interface. The line of the dislocation is into the page. Associated with the dislocation is an extra half plane of atoms in the SiC.

The overall structure of a dislocation can be quite complex. However, a dislocation can be understood to be a combination of the two basic types: *screw* and *edge* dislocations.

2.6.1 Screw Dislocations

A screw dislocation can be created in a regular crystal lattice by the application of a shear stress, as shown in Figure 2.18. Consider the plane ABCD, which is one of the regular planes of atoms in the crystal. Suppose a shear stress is applied to this plane, as shown schematically by the forces in the diagram. If this stress is sufficiently large (beyond the elastic limit for the crystal), it will cause the atoms on either side of the shear plane to be displaced by one atomic spacing. The line of the screw dislocation so formed is AB.

The arrangement of atoms around the screw dislocation forms a *single surface helicoid*, similar to a spiral staircase. Looking down the dislocation line AB, if the helix advances one plane for each clockwise rotation made around it, the dislocation is a *right-handed* screw dislocation. If the dislocation has the opposite sense, it is called a *left-handed* screw dislocation. The dislocation shown is therefore right-handed.

2.6.2 Edge Dislocations

An edge dislocation involves the inclusion of an extra half plane of atoms ABCD in an otherwise perfect crystal, as shown in Figure 2.19. Here the line of the dislocation AB is the edge of the extra half plane. Such a dislocation could be created by the application

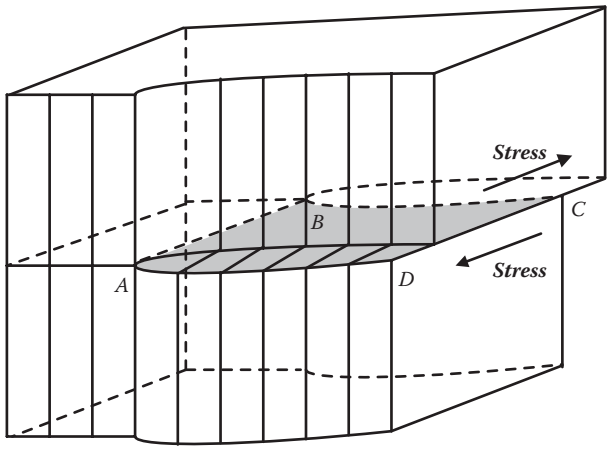


FIGURE 2.18
"Screw dislocation."

of a compressive stress as shown in the diagram. The edge dislocation shown is called a positive edge dislocation and is represented by the symbol ⊥ because the extra half plane of atoms has been added above the line AB. In a negative edge dislocation, the extra half plane would exist below the line AB.

2.6.3 Slip Systems

The geometry of a crystal dislocation is specified by its *line vector*, *Burgers vector*, and *glide plane*. The line vector **l** is a vector in the direction of the line of the dislocation. It need not

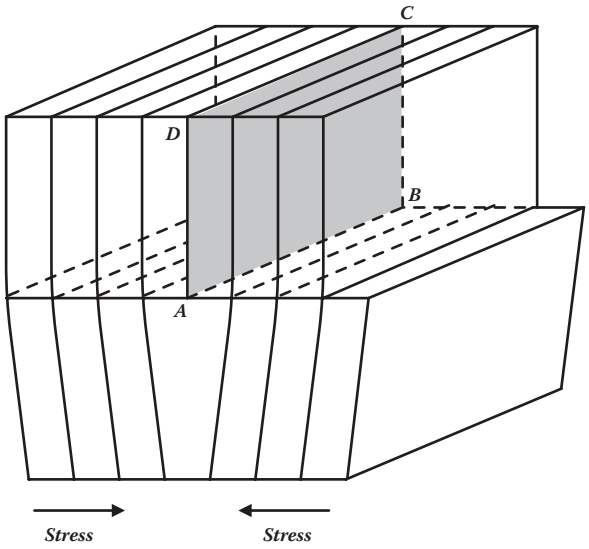


FIGURE 2.19
"Edge dislocation."

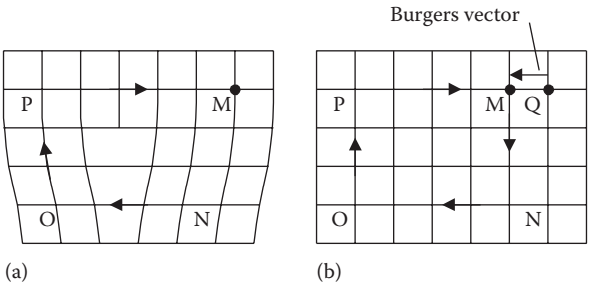


FIGURE 2.20
Burgers circuit. (a) The Burgers circuit MNOP starts and ends on the same point M and encloses a positive edge dislocation with its line into the plane of the paper. (b) In the perfect crystal, the same circuit starts at point M but fails to close, instead ending on point Q. The closure failure QM is the Burgers vector.

be a unit vector, and it is usually expressed as a basic lattice translation or combination of lattice translations. The Burgers vector may be determined by consideration of a *Burgers circuit*. A Burgers circuit is any atom-to-atom path that forms a closed loop around the dislocation core, in a clockwise direction, when looking in the line vector direction. For example, the path MNOPQ shown in Figure 2.20a is a Burgers circuit around an edge dislocation. (The line of the dislocation is into the plane of the page.) Suppose the same sequence of atom-to-atom jumps is made in a perfect crystal, as shown in Figure 2.20b. The failure of the Burgers circuit to close upon itself in the perfect crystal shows the presence of the dislocation, and the closure failure is the Burgers vector:

$$\mathbf{b} = \overrightarrow{QM} . \tag{2.152}$$

The character of a dislocation can be specified by the angle between the Burgers vector and the line vector. For an edge dislocation such as the one shown in Figure 2.20, the Burgers vector is always perpendicular to the line vector. Therefore, edge dislocations are sometimes referred to as 90° *dislocations*. For a screw dislocation, the line vector and Burgers vector are parallel, resulting in the terminology 0° *dislocation*. Although pure edge and screw dislocations are encountered in real crystals, dislocations of mixed character are far more common. For example, 60° misfit dislocations are most often observed in diamond and zinc blende crystals. The 60° dislocation exhibits an angle of 60° between the Burgers vector and the line vector. Its nature and core structure can therefore be considered part edge and part screw. In heteroepitaxial GaN on sapphire (0001), $a + c$ type dislocations exhibit mixed character whether they are located in the interface, as misfit dislocations, or along the c axis, as threading dislocations.[136] An example of a pure screw dislocation is a c-type threading dislocation, parallel to the c-axis, in heteroepitaxial GaN on sapphire (0001).[137] Pure edge dislocations have been observed at the substrate interface of GaSb/GaAs (001) grown in the "interfacial misfit (IMF) dislocation" mode.[138]

The Burgers vector is conserved for any dislocation passing through a crystal. Real dislocations are seldom perfectly straight but tend to follow paths with sometimes jagged changes in direction. Nonetheless, *any* Burgers circuit enclosing the dislocation will reveal the unique Burgers vector. The interesting implication is that any dislocation that changes direction changes character (the angle between the Burgers vector and the line vector changes along the dislocation). Therefore, a dislocation with screw character along part of its line may have 60° character or edge character elsewhere along its path.

For the determination of the Burgers vector, a clockwise path is taken around the Burgers circuit when looking down the line of the dislocation (in the direction of **l**). The Burgers vector is taken to run from the finish to the start of the Burgers circuit. This is the so-called *right-hand/finish–start (RH/FS) convention*. Note that the direction for the line vector can be arbitrarily assigned one of two ways. However, reversing the line vector also reverses the Burgers vector and preserves the angle between them.

The Burgers vector shows the direction and amount of slip associated with the crystal distortion that created the dislocation. Further distortion of the crystal in response to applied stresses may cause the dislocation to move by a mechanism called slip.* The *slip direction* is the same as the Burgers vector. Moreover, the *slip plane* is the plane containing the Burgers vector and the line vector.†

For a perfect, or unit, dislocation, the Burgers vector is a lattice translation vector. That is, the Burgers vector connects two lattice points in the perfect crystal. A perfect dislocation may dissociate into two partial dislocations, but the Burgers vector is conserved in the process. Thus, if a perfect dislocation with Burgers vector $\mathbf{b}_1$ dissociates into partial dislocations with Burgers vectors $\mathbf{b}_2$ and $\mathbf{b}_3$, then

$$\mathbf{b}_1 = \mathbf{b}_2 + \mathbf{b}_3 \,. \tag{2.153}$$

Reactions between two or more dislocations are possible as long as the total Burgers vector is conserved. An interesting special case is the reaction between two dislocations with opposite Burgers vector, which results in annihilation of both dislocations.

2.6.4 Dislocations in Diamond and Zinc Blende Crystals

The slip planes in a crystal are usually the planes with the highest density of atoms (the close-packed planes) because these have the greatest separation. In diamond and zinc blende semiconductors, the usual glide planes are the {111} planes. The direction of slip usually corresponds to the shortest lattice translation vector. Typically, slip directions (Burgers vectors) in the cubic semiconductors are of the type $a/2\langle 110\rangle$.

Cubic semiconductor crystals have four {111} planes with three ⟨110⟩ directions in each. Therefore, there are 12 distinct slip systems in a diamond or zinc blende crystal. Table 2.17 enumerates the 12 slip systems for a cubic semiconductor. A subset of these slip systems may be active during mismatched heteroepitaxy, depending on the crystal orientation. For example, eight of these are active for (001) heteroepitaxy.

The line vectors for dislocations in cubic semiconductors are typically of the type ⟨110⟩. Therefore dislocations on the 12 slip systems will be pure edge, pure screw, or 60° dislocations. All three types have been observed in heteroepitaxial zinc blende semiconductors, but 60° dislocations are the most prevalent. For the purpose of compact notation, the slip system with the Burgers vector $a/2[10\bar{1}]$ and the (111) glide plane would be called the $a/2[10\bar{1}](111)$ slip system. This class of slip systems would be referred to collectively as $a/2\langle 110\rangle\{111\}$.

Other types of slip systems have sometimes been observed in zinc blende semiconductors. V-shaped dislocations on $a/2\langle 110\rangle\{011\}$ slip systems have been found in

* Usually, motion of a single dislocation in this way is called "glide," and the term *slip* is usually used to mean the glide of many dislocations.

† For screw dislocations, the Burgers vector and line vector are parallel and share infinitely many planes. But real dislocations follow curved or jagged paths and change character along their path, eliminating this apparent ambiguity.

TABLE 2.17

Slip Systems in Diamond and Zinc Blende Crystals

Glide Plane	Burgers Vector
(111)	$\frac{a}{2}[10\bar{1}]$
(111)	$\frac{a}{2}[01\bar{1}]$
(111)	$\frac{a}{2}[1\bar{1}0]$
$(\bar{1}\bar{1}1)$	$\frac{a}{2}[101]$
$(\bar{1}\bar{1}1)$	$\frac{a}{2}[011]$
$(\bar{1}\bar{1}1)$	$\frac{a}{2}[1\bar{1}0]$
$(\bar{1}11)$	$\frac{a}{2}[101]$
$(\bar{1}11)$	$\frac{a}{2}[01\bar{1}]$
$(\bar{1}11)$	$\frac{a}{2}[110]$
$(1\bar{1}1)$	$\frac{a}{2}[10\bar{1}]$
$(1\bar{1}1)$	$\frac{a}{2}[011]$
$(1\bar{1}1)$	$\frac{a}{2}[110]$

heteroepitaxial layers. Chu and Nakahara[139] observed dislocations on a $a/2\langle 100\rangle\{100\}$ slip system in InGaAsP/InP (001). Cooman and Carter[140] reported dislocations on {100} slip planes in GaAs. In degraded zinc blende LD structures, numerous workers have identified dislocations on $\langle 110\rangle/\{111\}$ slip systems, which appear to be associated with the dark line defects (DLDs).[141]

Perfect dislocations in diamond or zinc blende semiconductors belong to either the glide or shuffle set. Consider the stacking sequence of (111) planes in either type of crystal. The stacking sequence is given by …AaBbCcAaBbCc…, as shown in Figure 2.21. A 60° dislocation of the shuffle set can be imagined as being constructed by making a cut at a shuffle plane, between planes of the same letter, followed by insertion of an extra half plane. A dislocation of the glide set can be constructed by a similar operation, with the cut made between different letter planes. Both types of dislocations are glissile. However, dislocations of the shuffle set have a line of interstitials or vacancies adjacent to their core. Movement of the row of point defects can occur only by shuffling, which greatly reduces the mobility of dislocations from the shuffle set. Following common practice in the field of heteroepitaxy, it will be assumed in this book that all dislocations are from the glide set.

In a zinc blende semiconductor, 60° dislocations of the glide set may be further classified as α and β dislocations according to the chemical makeup of their cores.[142,143] In the zinc blende semiconductor AB, the α dislocations have all A atoms at the core, whereas

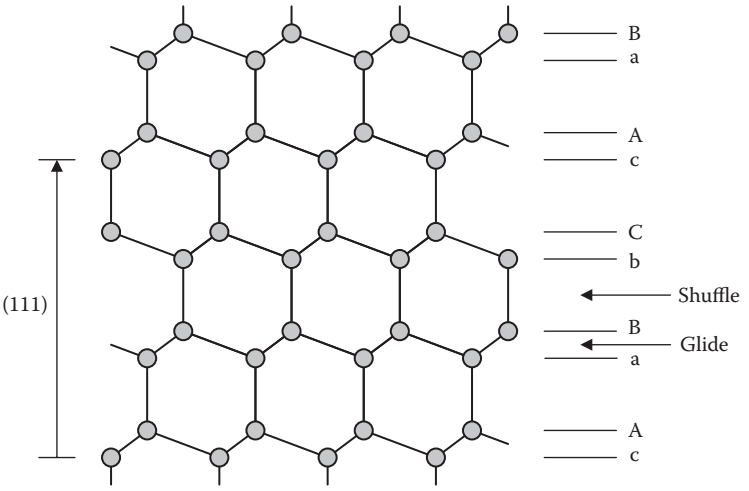


FIGURE 2.21
The stacking sequence for (111) planes in a diamond or zinc blende crystal ((01$\bar{1}$) projection). (Adapted from D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. [New York: Butterworth Heinemann, 2001]. Used with permission.)

β dislocations have all B atoms at their cores. α and β dislocations can be expected to behave differently due to their different core structures. Differences in mobility have been demonstrated for the two types of dislocations,[144] and differences in their dissociation to partial dislocations have also been shown.[145] These differences can be expected to affect the dynamics of lattice relaxation in mismatched heteroepitaxial layers. Chemical differences between the α and β dislocations have been demonstrated as well. Yokozeki et al.[146] investigated the passivation of misfit dislocations by elemental hydrogen in InAs/GaAs superlattices and found that β dislocations were more readily passivated due to the greater strength of the As-H bond compared with the Ga-H bond.

Note that in elemental semiconductors such as Si or Ge, the two sublattices have the same type of atoms, eliminating the distinction between α and β dislocations. The same also goes for SiGe alloys, in which the occupation of atomic sites is random and not ordered.

2.6.4.1 Threading Dislocations in Diamond and Zinc Blende Crystals

Threading dislocations of the edge, screw, and 60° types are present in bulk diamond and zinc blende crystals due to thermal and mechanical stresses acting on the crystal boules during growth or cooling. Some of these threading dislocations will intersect the surfaces of wafers cut from the crystal boules. A heteroepitaxial layer grown on such a wafer will typically inherit the threading dislocations from the substrate, which then propagate through the heteroepitaxial layer to a free surface.

The one-to-one relation between substrate and epitaxial layer dislocations has been established by a TEM study of epitaxial GaAs.[147] This study also showed that threading dislocations may cause one-to-*n* multiplication, whereby *n* threading dislocations propagate in the epitaxial layer. The importance of dislocation multiplication is also demonstrated by the observation in a number of mismatched heteroepitaxial systems that the epitaxial layers have dislocation densities orders of magnitude higher than the substrates.

The threading dislocation densities in semiconductor wafers vary greatly with the type of material. Three hundred millimeter silicon wafers[148] are virtually dislocation-free, with threading dislocation densities of $<10\,cm^{-2}$ (inferred from etch pit densities [EPDs]). For InP wafers,[149] the dislocation densities vary greatly depending on the wafer diameter and doping. Seventy-five millimeter iron-doped semi-insulating InP wafers exhibit EPDs of about $10^5\,cm^{-2}$, whereas 50 mm zinc-doped p-type wafers have EPDs of $<100\,cm^{-2}$. In the case of GaAs, 150 mm wafers[150] grown by the high-pressure liquid-encapsulated Czochralski process are available with dislocation densities of $<10^4\,cm^{-2}$. II-VI semiconductor substrates such as CdTe and ZnSe are available only in small sizes and tend to exhibit relatively high threading dislocation densities as a consequence of their low values for the critical resolved shear stress. In the case of bulk CdTe, the high density of dislocations tends to arrange in a subgrain structure.[150] However, bulk CdTe has been grown with EPDs of $<10^5\,cm^{-2}$ by a vapor growth process.[151]

2.6.4.2 Misfit Dislocations in Diamond and Zinc Blende Crystals

Misfit dislocations form at (or near) the interface to relieve strain in a mismatched heteroepitaxial layer, once this layer exceeds the critical layer thickness. In the case of (001) heteroepitaxy of cubic semiconductors, these misfit dislocations form along the two orthogonal ⟨110⟩ directions in the plane of the interface.[152] Figure 2.22 shows a rectangular array of misfit dislocations aligned with the ⟨110⟩ directions in a 20 nm thick layer of $In_{0.2}Ga_{0.8}As$/GaAs (001).

For heteroepitaxial layers with moderate mismatch ($|f|<1\%$), most of these misfit dislocations have 60° character, with Burgers vectors of the type $a/2\langle 101\rangle$. This is because any pure

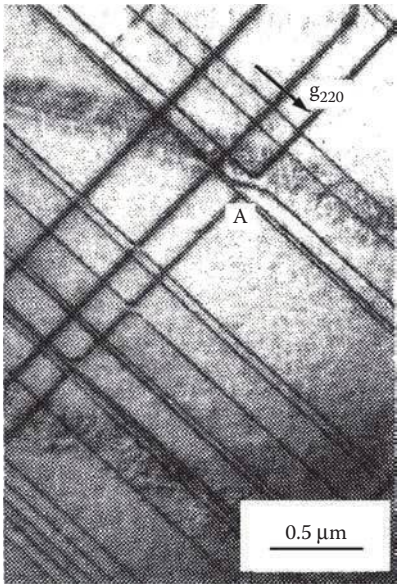


FIGURE 2.22
Plan-view TEM micrograph showing a rectangular array of misfit dislocations aligned with the directions at the interface of a 20 nm thick layer of In0.2Ga0.8As/GaAs (001). (Reprinted from R. H. Dixon and P. J. Goodhew, *J. Appl. Phys.*, 68, 3163 [1990]. With permission. Copyright 1990, American Institute of Physics.)

edge dislocation in the interface will have its Burgers vector (slip direction) in the interface as well. Such an edge dislocation is unable to glide from the free surface to the interface.

During the early stages of relaxation, misfit dislocations are not created in equal numbers along the two ⟨110⟩ directions,[153,154] possibly due to differences in mobility between the α and β dislocations. This is shown by the series of micrographs in Figure 2.23 for heteroepitaxial $In_{0.25}Ga_{0.75}As/GaAs$ (001). The thicknesses of the layers in Figure 2.23a–d are 20, 30, 40, and 60 nm, respectively. The average linear misfit dislocation density increases with thickness from 0.12×10^5 dislocations cm^{-1} for the 20 nm sample to 1.7×10^5 dislocations cm^{-1} for the 60 nm sample. However, the linear densities of misfit dislocations are different in the two orthogonal ⟨110⟩ directions, and this is most pronounced for the 30 nm sample of Figure 2.23b.

In highly mismatched heteroepitaxial layers, the misfit dislocation structure at the interface can be much less regular.[154] This is shown by the series of micrographs in Figure 2.24 for $In_xGa_{1-x}As/GaAs$ (001) samples of varying composition, and therefore mismatch. In the sample of Figure 2.24a, with $x=0.25$ and $f=-1.7\%$, the misfit dislocations are predominantly straight and aligned with the ⟨110⟩ directions. With increasing mismatch, however, the misfit dislocation structure becomes increasing irregular. The misfit dislocations in the sample shown in Figure 2.24c with $x=0.40$ and $f=-2.7\%$ exhibit a "puzzle-piece"

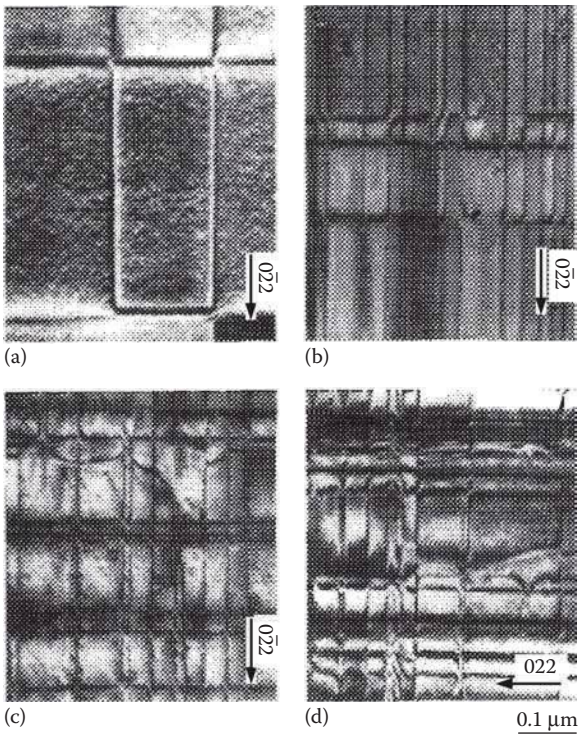


FIGURE 2.23
TEM micrographs showing the misfit dislocation structures in heteroepitaxial $In_{0.25}Ga_{0.75}As/GaAs$ (001) layers of increasing thickness. The mismatch is $f=-1.3\%$ and the critical layer thickness is $h_c=6$ nm. The actual thickness and linear density of misfit dislocations are (a) $h=20$ nm and $\rho=0.12\times10^5$ dislocations/cm, (b) $h=30$ nm and $\rho=0.9\times10^5$ dislocations/cm, (c) $h=40$ nm and $\rho=1.4\times10^5$ dislocations/cm, and (d) $h=60$ nm and $\rho=1.7\times10^5$ dislocations/cm. (Reprinted from K. R. Breen et al., *J. Vac. Sci. Technol. B*, 7, 758 [1989]. With permission. Copyright 1989, American Vacuum Society.)

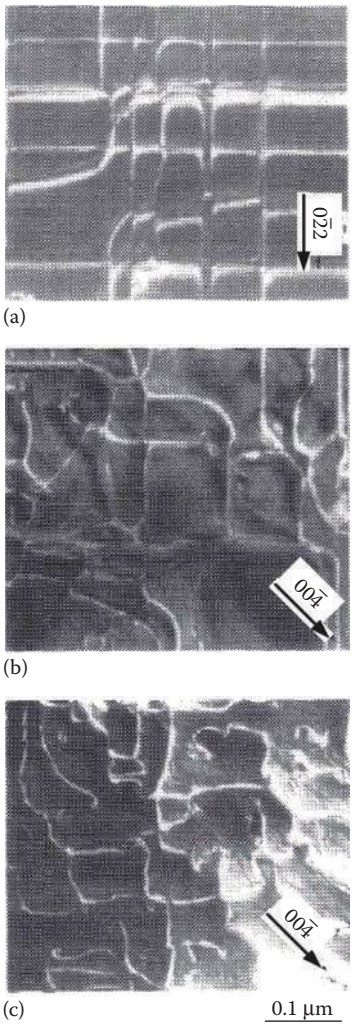


FIGURE 2.24
TEM micrographs showing the misfit dislocation structures in heteroepitaxial InxGa1-xAs/GaAs (001) layers of varying composition and mismatch. The composition, mismatch, and thickness are (a) $x=0.25$, $f=-1.7\%$, and $h=60\,\text{nm}$; (b) $x=30$, $f=-2\%$, and $h=30\,\text{nm}$; and (c) $x=0.40$, $f=-2.7\%$ and $h=20\,\text{nm}$. (Reprinted from K. R. Breen et al., *J. Vac. Sci. Technol. B*, 7, 758 [1989]. With permission. Copyright 1989, American Vacuum Society.)

structure characteristic of dislocations introduced following island growth of the highly mismatched heteroepitaxial layer.

Misfit dislocations with pure edge character can be created by the reaction of two 60° dislocations at the interface of a heteroepitaxial zinc blende semiconductor. For example, consider two 60° misfit dislocations, both with line vector [110], and with the Burgers vectors $a/2[101]$ and $a/2[0\bar{1}\bar{1}]$. They may combine to form a single edge dislocation by the reaction

$$\frac{a}{2}[101]+\frac{a}{2}\left[0\bar{1}\bar{1}\right]\rightarrow\frac{a}{2}\left[1\bar{1}0\right]. \tag{2.154}$$

Such a reaction is energetically favorable, as can be shown by Frank's rule. The resulting edge dislocation will relax as much mismatch strain as the two 60° dislocations, without introducing the tilt components of the Burgers vectors. The edge dislocations observed by TEM in moderately mismatched heteroepitaxial zinc blende semiconductors are believed to have formed by such a reaction mechanism. Although large numbers of pure edge dislocations are observed in the highly mismatched system GaSb/GaAs, the mechanism is not believed to be the reaction of 60° dislocations.[155] Interfacial screw dislocations are not expected to form by the reaction of 60° dislocations because they would have no mismatch-relieving component and are energetically unfavorable.

2.6.5 Dislocations in Wurtzite Crystals

In wurtzite semiconductors such as GaN, the common slip directions (Burgers vectors) are of the *a*-type $a/3\langle 11\bar{2}0\rangle$[156] and *c*-type $c\langle 0001\rangle$.[157] Mixed dislocations, with both *a*-type and *c*-type components, have also been observed.[158] The (0001) basal plane has the highest density of atoms, and so slip often occurs in this plane. This is referred to as basal slip. There is one basal plane with three $a/3\langle 11\bar{2}0\rangle$ slip directions, resulting in three basal slip systems.

Nonbasal slip is also possible and occurs on the first-order prism planes, of type $\{1\bar{1}00\}$, and on the first-order pyramidal planes, of type $\{10\bar{1}1\}$. Nonbasal slip systems can be important in the lattice relaxation process for (0001) heteroepitaxy of wurtzite semiconductors. This is because the slip planes intersect both the surface and the interface with the substrate, allowing dislocations to glide into the interface to relieve lattice mismatch strain.

2.6.5.1 Threading Dislocations in Wurtzite Crystals

Heteroepitaxial III-nitrides are typically grown on *c*-plane (0001) sapphire (α-Al_2O_3) substrates or *c*-plane (0001) 6H-SiC substrates. Low-temperature buffer layers of GaN[158] or AlN[159] are grown on the substrate prior to the deposition of device-quality layers. In either case, the initial growth mode is 3-D and a continuous heteroepitaxial film forms by the coalescence of islands. The structural evolution of such layers is rather complex, but some studies have indicated that the as-grown buffer (nucleation) layers are amorphous, and crystallize by solid phase epitaxy during a subsequent heat treatment. In one study of low-temperature GaN on *c*-plane sapphire it was found that the initial islands are zinc blende GaN[160] even though the final layer exhibits the wurtzite structure. In either case, it appears that the misfit and threading dislocations are introduced predominantly by injection at the edges of the islands formed during recrystallization, rather than by the bending over and multiplication of substrate threading dislocations, as in the heteroepitaxy of zinc blende crystals with moderate lattice mismatch.

The most common threading dislocations in these layers are pure edge dislocations, with [0001] line vectors and Burgers vectors of the *a*-type $a/3\langle 11\bar{2}0\rangle$; pure screw dislocations, with [0001] line vectors and Burgers vectors of the *c*-type $c\langle 0001\rangle$; and $a + c$ mixed dislocations with Burgers vectors of the type $a/3\langle 11\bar{2}3\rangle$.[163] The (nonbasal) slip for this system occurs on the first-order prism planes of the type $\{1\bar{1}00\}$.[161] Because the screw dislocations have a Burgers vector that is perpendicular to the interface, they are not associated with misfit dislocations in the *c*-plane interface. Instead, they may be introduced to relax the mismatch strain at the steps of vicinal (tilted) substrates.[162] It has been reported that the screw dislocations are suppressed by the use of a low-temperature AlN buffer layer on *c*-plane sapphire.

Heteroepitaxial layers of GaN grown on *c*-face SiC or sapphire substrates by metalorganic vapor phase epitaxy (MOVPE) typically contain a tangle of threading dislocations in the first 0.5 μm of thickness. Above this, there are relatively straight threading dislocations, aligned with the ⟨0001⟩ direction, with a relatively constant density of $10^8 - 10^9 cm^{-2}$. This general behavior is illustrated in Figure 2.25, which shows threading dislocations in a 7 μm thick layer of GaN grown on 6H-SiC (0001) with an AlN buffer by MOVPE.[163] Similar behavior is also observed in the case of GaN on sapphire (0001), as shown in Figure 2.26 for a 1.2 μm thick layer grown by MOVPE.

Knoke et al.[163] reported that in the low-pressure solution growth of GaN on (0001) sapphire the most common threading dislocations were *c*-type screw dislocations with a concentration of 4×10^9 cm^{-2}, followed by *a* + *c* type dislocations with a concentration of 1×10^9 cm^{-2}, and then *a*-type dislocations with a concentration of 0.1×10^9 cm^{-2}. However, they also reported that this material was facetted and that the threading dislocations bent toward facets and did not maintain ⟨0001⟩ line vectors.

Hollow-core threading dislocations, called nanopipes, are sometimes observed in thick GaN layers grown on sapphire by hydride vapor phase epitaxy (HVPE)[164] and have also

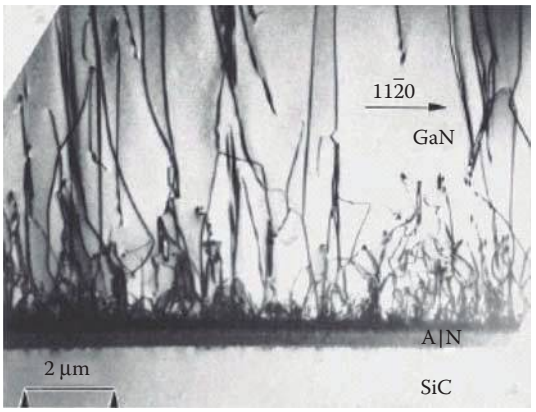


FIGURE 2.25
Cross-sectional TEM micrograph of GaN/AlN/6H-SiC (0001) showing the structure of the threading dislocations. (Reprinted from F. R. Chien et al., *Appl. Phys. Lett.*, 68, 2680 [1996]. With permission. Copyright 1996, American Institute of Physics.)

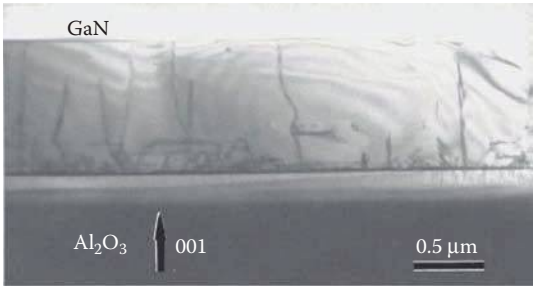


FIGURE 2.26
Cross-sectional TEM micrograph of GaN/α-Al_2O_3 (0001) showing the structure of the threading dislocations. (Reprinted from D. Kapolnek et al., *Appl. Phys. Lett.*, 67, 1541 [1995]. With permission. Copyright 1995, American Institute of Physics.)

been found in MOVPE-grown GaN/sapphire (0001).[165] These nanopipes have screw character and are fundamentally similar to the micropipes that are common in SiC. Compared with the micropipes in SiC, the GaN defects are found to have much smaller diameters (3.5–50 nm), as their name suggests.

2.6.5.2 Misfit Dislocations in III-Nitrides

In the case of (0001) heteroepitaxy of III-nitrides, misfit dislocations are introduced at the interface along $\langle 11\bar{2}0\rangle$ directions.[166] These misfit dislocations usually have edge character and represent the terminations of $\{11\bar{2}0\}$ planes. There are three equivalent $\langle 11\bar{2}0\rangle$ directions in the basal plane, and the misfit dislocations would meet each other at 60° angles if they formed a regular, triangular array.

Plan-view TEM micrographs show that, in the case of (almost completely relaxed) AlN/sapphire (0001), with $f = -13\%$, the misfit dislocations are evenly spaced along a $\langle 1\bar{1}00\rangle$ direction with a spacing of 2.0 nm.[167] However, plan-view TEM micrographs of GaN/sapphire (0001), with $f = -17\%$, reveal an irregular, puzzle-piece structure,[167] as shown in Figure 2.27. This is expected for any highly mismatched heteroepitaxial layer with a 3-D growth mode. It is possible that the misfit dislocations at a low-mismatch interface between III-nitrides will assume a regular, triangular pattern, but such results have not yet been reported.

2.6.6 Dislocations in Hexagonal SiC

Micropipes[168] are defects unique to SiC, which degrade the performance of devices. These defects are actually the hollow cores of axial screw dislocations with [0001] line vectors. The Burgers vectors of these hollow-core dislocations are multiples of the *c* lattice parameter.[169] In both the 4H and 6H polytypes, screw dislocations with Burgers vectors of *c*[0001] or *c*[0002] have closed cores. Super screw dislocations with Burgers vectors of *c*[0004] or greater invariably have hollow cores, or micropipes, associated with them, and their existence is driven by line energy considerations. The 3c screw dislocations, with Burgers vectors of *c*[0003], have been observed with hollow and closed cores. The dissociation of super screw dislocations into multiple screw dislocations can therefore eliminate micropipes if the resulting dislocations have closed cores.[170,171] This process is called micropipe sealing.

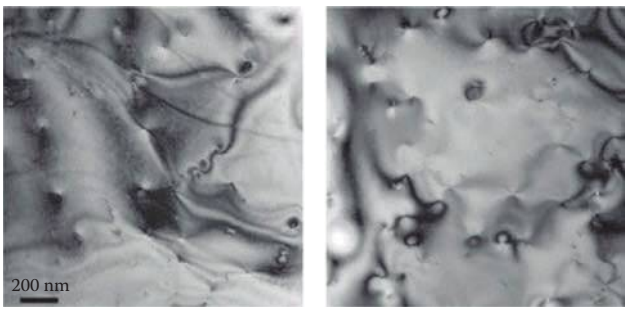


FIGURE 2.27
Plan-view TEM micrographs showing the misfit dislocations in 0.6 μm thick GaN/sapphire (0001) grown by MOVPE. (a) V/III ratio = 2100 and dislocation density = 2.4×10^9 cm^{-2}. (b) V/III ratio = 2600 and dislocation density = 3.6×10^9 cm^{-2}. (Reprinted from H. P. D. Schenk et al., *J. Cryst. Growth*, 258, 232 [2003]. With permission. Copyright 2003, Elsevier.)

As with wurtzite semiconductors, we would also expect dislocations with Burgers vectors of the type $a/3\langle 11\bar{2}0\rangle$; indeed, pure edge dislocations of this type have been observed in bulk 4H-SiC crystals,[172] with line vectors along [0001]. Molten KOH etching studies have revealed that these threading edge defects sometimes exist in linear arrays, which constitute low-angle grain boundaries[173,174] (also called domain walls).

2.6.6.1 Threading Dislocations in Hexagonal SiC

SiC device layers are often grown homoepitaxially on commercially available 4H- and 6H-SiC wafers. These substrates contain 10^3-10^4 cm^{-2} screw dislocations and 10^4-10^5 cm^{-2} edge dislocations,[175,176] both of which run parallel to the [0001]. As a consequence, they emerge at the surface and will replicate in epitaxial layers. Basal plane dislocations (screw dislocations lying within the (0001) plane) are also present, and emerge at the top surface in vicinal substrates, which are typically miscut by 8° from the (0001). In addition, wafers of either polytype typically contain $10-100$ cm^{-2} micropipes, which thread to the surface. In recent years, 4H-SiC wafers with "ultra-low micropipe densities" of less than 5 cm^{-2} have become available.[177]

2.6.7 Strain Fields and Line Energies of Dislocations

A dislocation line is surrounded by a strain field, which raises the energy of the crystal and also interacts with externally applied stresses. The elastic strain energy is also an important contribution to the dislocation line energy, with the balance determined by the inelastic contribution of the dislocation core. Strain field interactions give rise to dislocation motion in stressed crystals, and also cause pairs of dislocations to repel or attract.

In this section, the strain fields will be given for pure screw and pure edge dislocations, with the simplifying assumption of an isotropic crystal. Using these, the line energies will be derived for the basic types of dislocations by integration of the elastic strain energy. The total line energy will be estimated by assigning a core energy term to the elastic line energy. The line energy of a mixed dislocation will be calculated using the above framework and the principle of superposition. These calculations represent reasonable approximations for cubic materials, but fail to provide a good description for nitride semiconductors, which will be considered separately.

2.6.7.1 Energy of the Screw Dislocation

Consider a straight screw dislocation lying along the z axis in a right-handed Cartesian coordinate system with $\mathbf{l}=[001]$ and $\mathbf{b}=[00b]$, as shown in Figure 2.28. In the case of an isotropic crystal, the strain field surrounding the screw dislocation is given by[178]

$$\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon_{zz}=\varepsilon_{xy}=\varepsilon_{yx}=0, \tag{2.155}$$

$$\varepsilon_{xz}=\varepsilon_{zx}=-\frac{by}{4\pi\left(x^2+y^2\right)}, \tag{2.156}$$

and

$$\varepsilon_{yz}=\varepsilon_{zy}=\frac{bx}{4\pi\left(x^2+y^2\right)}. \tag{2.157}$$

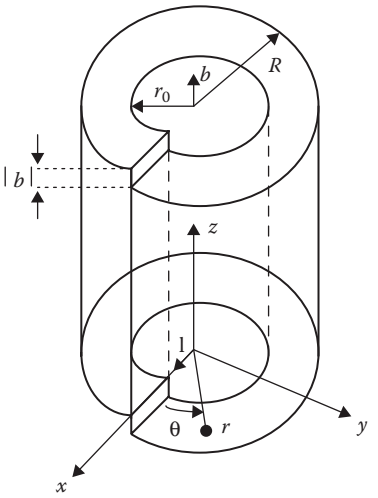


FIGURE 2.28
Model of screw dislocation for calculation of the strain energy.

Thus, only shear strains are associated with the screw dislocation. Transforming to cylindrical coordinates, we have

$$\varepsilon_{\theta z} = \varepsilon_{z\theta} = \frac{b}{4\pi r}, \tag{2.158}$$

and from Hooke's law,

$$\sigma_{\theta z} = \sigma_{z\theta} = \frac{Gb}{2\pi r}. \tag{2.159}$$

The bulk elastic strain energy per unit length of dislocation may be calculated by integrating Equation 2.159 over the plane perpendicular to the dislocation. Note that the shear strains are proportional to $1/r$ and approach infinity as $r \to 0$. However, the strains exceed the limits of linear elasticity theory (Hooke's law) in the core of the dislocation. To avoid this difficulty, we set the lower limit of integration at r_0, which is the radius of the dislocation core. Then, for the screw dislocation in an isotropic crystal,

$$E_{el}(screw) = \int_{r_0}^{R} 4\pi r dr G \varepsilon_{\theta z}^2 = \frac{Gb}{4\pi}\int_{r_0}^{R}\frac{dr}{r} = \frac{Gb^2}{4\pi}\ln\left(\frac{R}{r_0}\right). \tag{2.160}$$

Calculation of the total energy per unit length of dislocation requires adding a core energy, which in general will include non-Hookian elastic energies, as well as the energy of dangling bonds. Two approaches to including the core energy are as follows: (1) the core energy term may be accounted for by adjusting the cutoff parameter r_0 to some value much less than b, the length of the Burgers vector,[179] or (2) a value of b is assigned somewhat arbitrarily to the cutoff parameter r_0, and a "core parameter" is added to the logarithm in Equation 2.160.[180] The latter approach is often used in semiconductor work and will be adopted here as well.

Estimates of the core energy are necessarily very approximate because the core structure is complex and poorly understood. A discrete elasticity theory has been applied to screw dislocations in alkali-halide crystals,[181] and the extension of this analysis yields an

estimate of 1.4 for the core parameter in diamond-type semiconductors.[182] Other estimates of the core parameter for screw and edge dislocations range from 1 to 2 for cubic semiconductors[180,183] and GaN with the wurtzite structure.[184,185] However, it is likely that these values are still overestimates because they assume linear elastic behavior for atomic displacements of up to one-half the relaxed atomic spacing.[182] A core parameter of 1 will be commonly used in this text. Therefore, the line energy of the screw dislocation is

$$E_{el}(screw) = \frac{Gb^2}{4\pi}\left[\ln\left(\frac{R}{b}\right)+1\right]. \tag{2.161}$$

Often, the logarithmic term in Equation 2.161 is modified as $\ln(\beta R/b)$, where β is a semiempirical dislocation core constant. Equation 2.61 could more generally written as

$$E_{el}(screw) = \frac{Gb^2}{4\pi}\ln\left(\beta\frac{R}{b}\right). \tag{2.162}$$

A dislocation core constant of $\beta = e$ corresponds to the dislocation core parameter of 1 used in this text. There is a line tension that resists the lengthening of a dislocation due to its finite line energy. This line tension is just equal to the line energy. The reason is that the work W done to lengthen a dislocation by an amount dl is the line tension times this length, $W = Fdl$, and $F = W/dl$ is the line energy. In the literature, the terms *line energy* and *line tension* are often used interchangeably.

2.6.7.2 Energy of the Edge Dislocation

Now consider a straight edge dislocation lying along the z axis in a right-handed Cartesian coordinate system with $\mathbf{l} = [001]$ and $\mathbf{b} = [\bar{b}00]$, as shown in Figure 2.29. In the case of an isotropic crystal, the strain field surrounding the edge dislocation is given by[179]

$$\varepsilon_{zz} = \varepsilon_{xz} = \varepsilon_{zx} = \varepsilon_{zy} = \varepsilon_{yz} = 0, \tag{2.163}$$

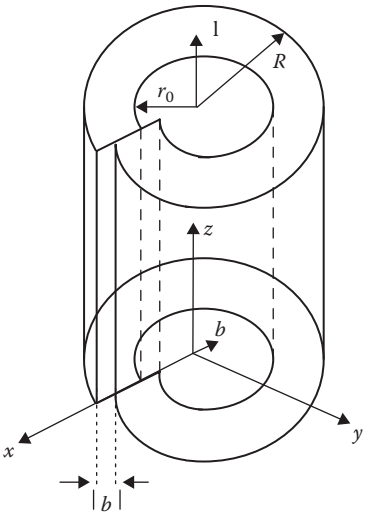


FIGURE 2.29
Model of edge dislocation for calculation of the strain energy.

$$\varepsilon_{xx} = \frac{by\left(x^2 - y^2\right)}{4\pi(1-\nu)\left(x^2+y^2\right)^2} - \frac{by}{2\pi\left(x^2+y^2\right)}, \tag{2.164}$$

$$\varepsilon_{xy} = \varepsilon_{yx} = \frac{bx\left(x^2 - y^2\right)}{4\pi(1-\nu)\left(x^2+y^2\right)^2}, \tag{2.165}$$

and

$$\varepsilon_{yy} = \frac{by\left(3x^2 + y^2\right)}{4\pi(1-\nu)\left(x^2+y^2\right)^2} - \frac{by}{2\pi\left(x^2+y^2\right)}, \tag{2.166}$$

where ν is the Poisson ratio. By a transformation to cylindrical polar coordinates it can be shown that the elastic line energy of the edge dislocation is

$$E_{el}(\text{edge}) = \frac{Gb^2}{4\pi(1-\nu)} \ln\left(\frac{R}{r_0}\right). \tag{2.167}$$

In general, utilizing the dislocation core constant β, as in the case of the screw dislocation, and assuming $r_0 = b$, the total line energy per unit length for the edge dislocation is

$$E_{el}(\text{edge}) = \frac{Gb^2}{4\pi(1-\nu)} \ln\left(\beta\frac{R}{b}\right). \tag{2.168}$$

Adopting a core parameter of 1 as in the case of the screw dislocation, the total line energy per unit length for the edge dislocation is modified to

$$E_{el}(\text{edge}) = \frac{Gb^2}{4\pi(1-\nu)}\left[\ln\left(\frac{R}{b}\right)+1\right]. \tag{2.169}$$

2.6.7.3 Energy of Mixed Dislocations

For a dislocation of mixed character, the strain field is the superposition of the individual strain fields for its edge and screw components. There is no interaction between the two component strain fields, so the line energy is the sum of the screw and edge contributions,

$$E_{el}(\text{mixed}) = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)} \ln\left(\beta\frac{R}{b}\right), \tag{2.170}$$

where α is the angle between the Burgers vector and the line vector. Adopting a dislocation core parameter of 1, Equation 2.170 is modified to

$$E(\text{mixed}) = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{R}{b}\right)+1\right]. \tag{2.171}$$

2.6.7.4 Frank's Rule for Dislocation Energies

The line energy of a dislocation has a relatively weak dependence on the dislocation character, and for any dislocation,

$$E \approx CGb^2 , \tag{2.172}$$

with $0.5 < C < 1$.[186] This is the basis for Frank's rule for dislocation reactions: a dislocation reaction is energetically favorable if $\sum b^2$ for the product(s) is less than $\sum b^2$ for the reactant(s). For example, two dislocations with opposite Burgers vectors may react and annihilate. In this case, $\sum b^2 = 0$ for the final configuration, so the annihilation reaction is energetically favored. Frank's rule can also be applied to the dissociation of perfect dislocations into partial dislocations.

2.6.7.5 Line Energies of Dislocations in Nitride Semiconductors

In the case of nitride semiconductors, the isotropic approximation is often inadequate. Therefore, the dislocation line energy depends on its line direction, as well as the Burgers vector. Following the approach of Steeds,[187] Holec et al.[188] have calculated the elastic contribution to the dislocation line energy for the case of a straight dislocation lying on the (0001) plane. The total dislocation energy per unit length includes the elastic and core terms and is given by

$$E_d = \left(Ab_{e,||}^2 + Bb_{e,\perp}^2 + Cb_s^2 \right) \ln\left(\frac{R}{r_c} \right) + E_{core} , \tag{2.173}$$

where A, B, and C are material constants related to the elastic stiffness coefficients, and are given in Figure 2.30 for the case of $In_xGa_{1-x}N$; $b_{e,||}$ and $b_{e,\perp}$ are the in-plane and out-of-plane edge components of the Burgers vector; b_s is the screw component of the Burgers

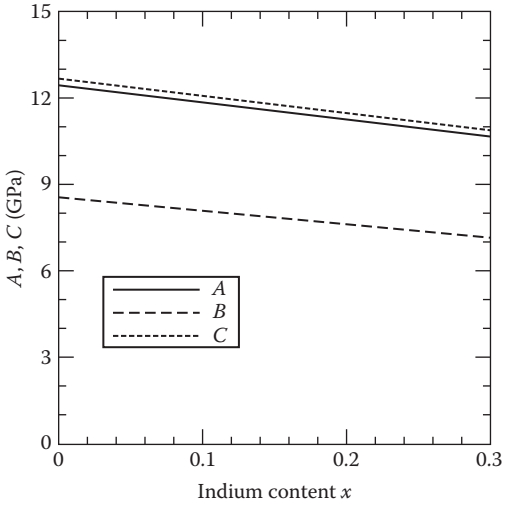


FIGURE 2.30
Values of the coefficients in the dislocation line energy calculation, for $In_xGa_{1-x}N$. (Reprinted from D. Holec et al., *J. Cryst. Growth*, 303, 314 [2007]. With permission. Copyright 2007, Elsevier.)

TABLE 2.18

Parameters of the 5/7-Atom Ring Dislocation Cores in Nitride Semiconductors as Reported by Belabbas et al.,[189] Lymperakis,[190] Kioseoglou et al.,[191] and Holec et al.[188]

Material	Dislocation	r_c (nm)	E_{core} (eV/Å)
GaN	$(a + c)$ type	0.72	3.12
GaN	a type	0.60	1.61
AlN	a type	0.83	1.71
InN	a type	0.54	1.66

vector; R is the film thickness or dislocation separation, whichever is smaller; and r_c is the cutoff radius for the dislocation core. In some work, it was assumed that $r_c \approx b/2$, but based on the calculations of Belabbas et al.,[189] Holec et al. assumed $r_c \approx b$ for $a+c$ type dislocations, and based on the calculations of Lymperakis[190] and Kioseoglou et al.,[191] Holec et al. assumed $r_c \approx 2b$ for a-type dislocations. Values for the dislocation core parameters are given in Table 2.18 for binary nitride semiconductors.

2.6.7.6 Hollow-Core Dislocations (Micropipes)

Because the line energy of a closed dislocation increases with the square of the Burgers vector length, dislocations with very long Burgers vectors sometimes develop hollow cores. This reduces the elastic strain energy associated with the dislocation, while introducing the surface energy associated with the hollow tube. Frank first derived an expression for the diameter of such a hollow core,[192] based on minimization of the free energy. Thus, in equilibrium the increase of free energy due to the addition of surface area should equal the free energy released by expanding the hollow core of the dislocation. If r is the radius of the hollow core, then the change in free energy per unit length of dislocation associated with a change in radius dr is given by

$$dU \approx \gamma 2\pi dr - \frac{Gb^2}{4\pi r} dr = 0, \tag{2.174}$$

where γ is the surface energy of the crystal and G is the shear modulus. Here, the crystal was assumed to be isotropic and the dislocation core parameter was neglected. Solving, we find the equilibrium core radius is given by

$$r = \frac{Gb^2}{8\pi^2\gamma}. \tag{2.175}$$

Hollow-core dislocations are common in the 4H, 6H, and 15R polytypes of SiC. These micropipes are screw dislocations with Burgers vectors aligned with the sixfold axis and having lengths of c, 2c, 3c, and so on. In 4H-SiC, the experimentally determined diameters for hollow cores can be fit[193] with the assumption of $\gamma/G = 1.2 \times 10^{-12}$ m. Thus, 3c super screw dislocations have been observed to have hollow cores approximately 180 nm in diameter, and 4c super screw dislocations have hollow cores of approximately 320 nm. For 6H-SiC, the c parameter is longer, and so the corresponding core diameters are expected to be 400 and 710 nm, respectively. It has been reported that screw dislocations having Burgers

vectors of length c, 2c, or 3c in 4H- or 6H-SiC need not have open cores, but those with 4c or greater are always reported to be open.

2.6.8 Forces on Dislocations

The dislocations in a crystal will move under the influence of an applied stress. The glide of many dislocations results in slip, which is manifested as a successive displacement of many planes of atoms along a slip plane, as shown in Figure 2.31a. Figure 2.31b shows a misfit dislocation along with its associated threading dislocation, and their orientation with respect to the slip plane. The load producing the applied stress therefore does work on the crystal, and the dislocation therefore responds as though it experiences a force equal to the work done divided by the distance moved.[194]

The force on a dislocation in a crystal with an arbitrary stress tensor σ is given by the Peach–Koehler formula:[195]

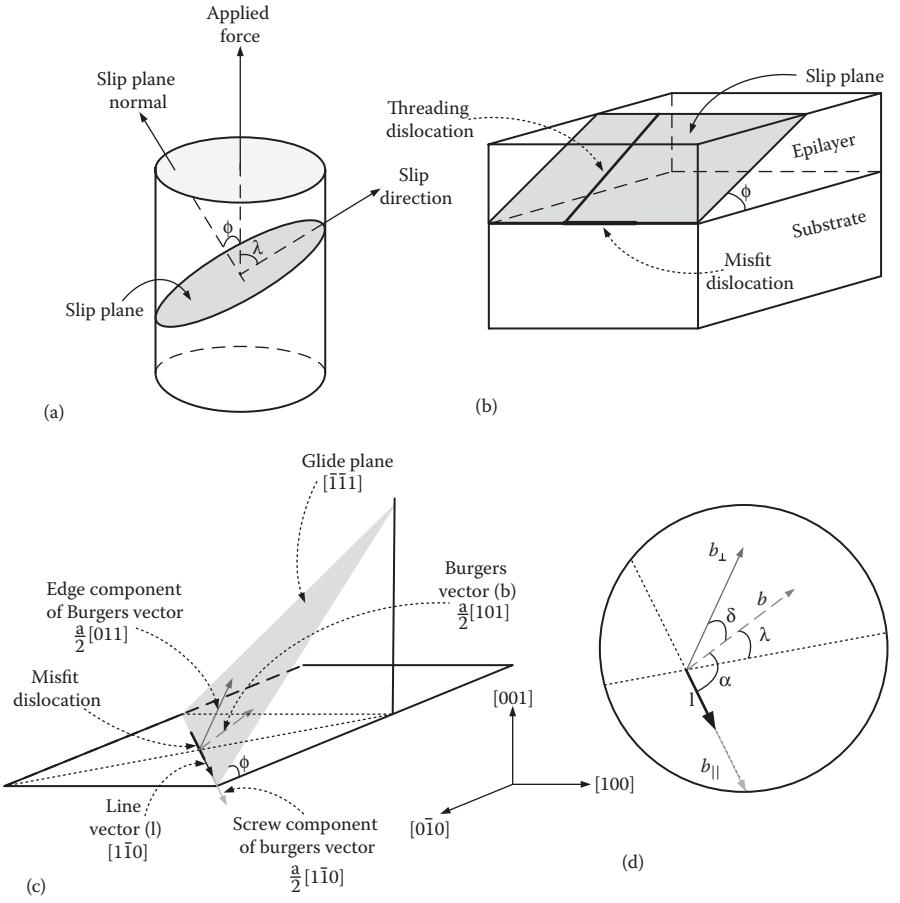


FIGURE 2.31
(a) Schematic representation of slip in crystalline material. (b) Interfacial misfit and threading dislocation orientation in a slip plane. (c) The geometrical components associated with the Burgers vector. (d) Angles between the Burgers vector and line vector, edge and screw components.

$$\frac{\mathrm{F}}{L} = (\sigma \cdot \mathbf{b}) \times \mathbf{s}, \tag{2.176}$$

where F/*L* is the vector force per unit length, *b* is the Burgers vector, and *s* is the unit vector in the direction of the line of the vector. In scalar form,

$$F = \tau b, \tag{2.177}$$

where *b* is the length of the Burgers vector and τ is the *shear stress, resolved on the slip plane, in the slip direction*. If the stress* in the crystal is produced by a tensile force *F* applied to a cross section of area *A*, then the stress is $\sigma = F/A$ and the resolved shear stress is

$$\tau_{res} = \sigma \cos\phi \cos\lambda, \tag{2.178}$$

where ϕ is the angle between the applied force and the normal to the slip plane, and λ is the angle between the applied force and the slip direction, as shown in Figure 2.31c and d. The quantity $\cos\phi\cos\lambda$ is referred to as the *Schmid factor*.[196] The most common misfit dislocations generated in the zinc blende epilayer are 60° in character and contain a Burgers vector of the type <110>; these vectors lie at 60° to the <110> line directions and 45° to the interfacial plane. Furthermore, because these are considered mixed dislocations, each 60° dislocation contains a screw and edge component (Figure 2.31c). However, only the edge component of the Burgers vector (the component that lies in the interfacial plane) acts in relaxing the mismatch strain. Therefore, based on the geometry of Figure 2.31c and d, the edge component of the Burgers vector is given by

$$b_{\perp} = b \sin\alpha \cdot \tag{2.179}$$

2.6.9 Dislocation Motion

Dislocations move by glide, climb, or a combination of both. Glide is motion in the direction of the Burgers vector and is called conservative motion. Climb is motion out of the glide plane (nonconservative motion). Both processes are thermally activated because they involve the breaking of crystal bonds, but climb requires long-range diffusion and is only important at very high temperatures.

Dislocation glide velocities have been determined by the "double-etch" technique in a number of crystals.[197] This method involves the use of a crystallographic etch before and after stressing the crystal. The etch will reveal sharp bottom pits at the places where dislocations emerge at the surface. If a dislocation moves while the crystal is being stressed, a new, sharp-bottomed pit will be produced at the new point of emergence for the dislocation. At the same time, the original pit will enlarge and take on a flat bottom as a result of the additional etching. Multiple cycles of stressing and etching can therefore be used to track the motion of individual dislocations. This enables basic studies of dislocation motion, which reveal the dependence of the dislocation glide velocities on temperature and applied stress.

Dislocation glide velocities have been measured by the double-etch method in GaAs,[198,199] Ge,[200,201] Si,[202] and InSb,[201,202] but also by direct observation in Ge_xSi_{1-x}/Si using TEM.[202,203] In these materials it has been found that the stress and temperature dependence of the dislocation glide velocity may be described by a phenomenological relationship,[204]

* In published work, the stress is usually denoted by two different symbols τ and σ, which both mean the same thing: In the descriptions which follow, we will generally use the notation τ for consistency.

$$v = v_0 \exp\left(-\frac{\tau_0}{\tau_{res}}\right)\exp\left(-\frac{U}{k_B T}\right), \tag{2.180}$$

where v_0 is a characteristic velocity, τ_0 is a characteristic stress, τ_{res} is the resolved shear stress on the glide plane in the slip direction, U is the activation energy for glide, T is the absolute temperature, and k_B is the Boltzmann constant. Over restricted ranges of stress, this relationship is commonly approximated by

$$v = v_0 \tau^m \exp\left(-\frac{U}{k_B T}\right), \tag{2.181}$$

where m and B are constants.

In bulk GaAs, it has been found experimentally that both m and U depend on the conductivity type (n-type or p-type), as well as the dislocation character.[199] Dislocation glide has been found to be affected by isoelectronic impurities,[67,205] as well as dopants and complexes involving dopant atoms interacting with point defects such as vacancies.[206] The values reported in the literature are in the range $1 < m < 3$, and often it is assumed that $m = 1$. It should be emphasized that Equation 2.181 is empirical in nature, and that it has been reported that the activation energy U is stress dependent[207] in SiGe alloys. For example, Hull et al.[203] modeled the dislocation velocities in $Si_{0.75}Ge_{0.25}$/Si (001) by

$$v = v_0 \tau_{\text{eff}} \exp\left(-\frac{E_v}{k_B T}\right), \tag{2.182}$$

where τ_{eff} is the effective stress* required to overcome the line tension and start the slip. $v_0 = 3.3 \times 10^{-7}\,\text{cm}^3\,\text{s}^{-1}\,\text{dyn}^{-1}$, and the activation energy was found to be stress dependent,

$$E_v = 2.0eV\left[1 - \frac{\tau_{\text{eff}}}{1.5 \times 10^{10}\,\text{dyn/cm}^2}\right]. \tag{2.183}$$

An alternate model for dislocation glide is the kink model described by Hirth and Lothe.[208] This model is physics based rather than phenomenological and has been applied to strained SiGe/Si (001) layers by Hull et al.[209] According to this model, the mechanism of dislocation glide involves the nucleation and propagation of kinks along the dislocation. This is illustrated in Figure 2.32 for the cases of (a) double kinks and (b) single kinks.

The nucleation rate for double kinks is given by

$$J = \frac{v_D b q \sigma_{\text{eff}}}{k_B T} \exp\left(-\frac{2F_k}{k_B T}\right), \tag{2.184}$$

where v_D is the attempt frequency, assumed to be the Debye frequency, which is $1.3 \times 10^{13}\,\text{s}^{-1}$ for Si; b is the length of the Burgers vector; q is the distance between Peierl's valleys, which is 0.33 nm in Si; $2F_k$ is the energy required to nucleate a double kink; k_B is the Boltzmann

* In published work, the effective stress (or excess stress) has been denoted by four different symbols, which all mean the same thing: τ_{eff}, τ_{exc}, σ_{eff}, and σ_{exc}. In the descriptions which follow, we will generally use the notation τ_{eff} for consistency.

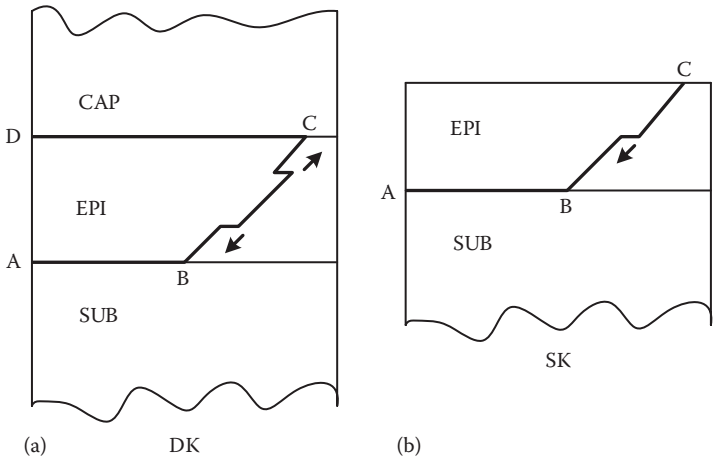


FIGURE 2.32
Diffusive kink model for the glide of a dislocation. (a) Double-kink model. (b) Single-kink model. (Reprinted from R. Hull et al., *J. Appl. Phys.*, 70, 2052 [1991]. With permission. Copyright 1991, American Institute of Physics.)

constant; T is the absolute temperature; and σ_{exc} is the Dodson and Tsao excess stress in the film, given by

$$\sigma_{\text{eff}} = \frac{2G(1+\nu)}{(1-\nu)}\left(\varepsilon_{||} - \varepsilon_{eq}\right), \tag{2.185}$$

where G is the shear modulus, ν is the Poisson ratio, $\varepsilon_{||}$ is the in-plane strain, and ε_{eq} is the equilibrium in-plane strain.

The kink propagation velocity is

$$v_k = \frac{v_D a^2 bq\sigma_{\text{eff}}}{k_B T}\exp\left(-\frac{E_m}{k_B T}\right), \tag{2.186}$$

where a is the kink jump distance, 0.38 nm in Si, and E_m is the energy barrier for kink jumps along the line direction. Due to kink–kink interactions in the presence of excess stress, Seeger and Schiller[210] showed that the double-kink nucleation energy is modified to

$$2F_k' = 2F_k - \sqrt{\frac{G(1+\nu)\sigma_{\text{eff}} b^3 q^3}{2\pi(1-\nu)}}\,. \tag{2.187}$$

Accounting for both the nucleation and propagation of kinks, the transverse velocity for a gliding threading dislocation is

$$v = \frac{2v_D abq^2\sigma_{\text{eff}}}{k_B T}\left(\frac{h}{h+X}\right)\exp\left(-\frac{E_m + F_k'}{k_B T}\right), \tag{2.188}$$

where h is the layer thickness (length of the gliding dislocation) and X is the average separation between double kinks, given by

$$X = 2a \exp\left(\frac{F_k'}{k_B T}\right). \tag{2.189}$$

There are two limiting cases for the dislocation glide velocity in the diffusive kink model. In thin, capped layers for which $h \ll L$, where L is the length of the propagating dislocation segment the double-kink model applies and

$$v = \frac{v_D b q^2 \sigma_{\text{eff}} h}{k_B T} \exp\left(-\frac{E_m + 2F_k'}{k_B T}\right) \tag{2.190}$$

In the bulk limit ($h \gg L$), or for uncapped layers, the single-kink model applies and

$$v = \frac{2 v_D a b q^2 \sigma_{\text{eff}}}{k_B T} \exp\left(-\frac{E_m + F_k'}{k_B T}\right) \tag{2.191}$$

The diffusive kink model has been applied by Hull et al. to strained SiGe/Si (001) layers with good agreement to experimental results. However, in that work the authors partitioned F_k and E_m in order to fit the data. It is known that for bulk Si, $F_k + E_m = 2.2\,\text{eV}$, and for bulk Ge, $F_k + E_m = 1.6\,\text{eV}$. Also, it has been found that $F_k = 1.0\,\text{eV}$ in bulk Si, but it is not known how the total splits between the two contributions for Ge or SiGe alloys. Therefore, Hull et al. used the value of F_k as a fitting parameter in their work. Even so, they found that the diffusive kink model overestimated the dislocation velocities for large values of excess stress (~1 MPa), and they attributed this to an overestimation of the Seeger and Schiller reduction in the activation energy at high values of excess stress.

2.6.10 Electronic Properties of Dislocations

The electronic properties of dislocations vary greatly among the different classes of semiconductor crystals. In diamond and zinc blende crystals, dislocations have been shown to be detrimental to device performance; in some cases, the presence of a single threading dislocation can lead to failure in a LD. The III-nitrides seem relatively immune to these same effects, so that LEDs fabricated in material with high dislocation densities exhibit little degradation and long lifetimes. In the common polytypes of SiC, the most important defects are micropipes (hollow-core, super screw dislocations), which cause failure of high-voltage devices.

2.6.10.1 Diamond and Zinc Blende Semiconductors

Dislocations have been found to act as nonradiative recombination centers in the arsenides,[211] phosphides,[212] nitrides,[213] and II-VI semiconductors.[214] In a spatially resolved photoluminescence study of dislocations in GaAs and InP,[215] it was found that the overall photoluminescence intensity was reduced significantly in a region of 5–10 μm around dislocations. This has been attributed to a reduction of the minority carrier lifetime in the vicinity of the dislocation. This may be due to increased rates of nonradiative bulk recombination mechanisms, as well as increased concentrations of lifetime-killing point defects segregated to the dislocations.

The minority carrier lifetime in dislocated GaAs has been modeled by Yamaguchi and Amano[216] with the assumption that the dislocations act as infinite sinks for minority carriers. In this model, the one-dimensional continuity equation for the transport of minority carriers to the dislocation is solved with the boundary condition that the minority carrier concentration is zero at the dislocation core. The solution is[217]

$$\frac{1}{\tau_d} = \frac{\pi^2 \mu k_B T D}{4q}, \tag{2.192}$$

where τ_d is the minority carrier lifetime associated with recombination mediated by dislocations, D is the dislocation density, μ is the mobility of minority carriers, k_B is the Boltzmann constant, T is the absolute temperature, and q is the electronic charge. The overall minority carrier lifetime τ is then given by

$$\frac{1}{\tau} = \frac{1}{\tau_0} + \frac{1}{\tau_d}, \tag{2.193}$$

where τ_0 is the lifetime associated with minority carrier recombination in the dislocation-free material, associated with point defects and intrinsic recombination processes.

Figure 2.33 shows the minority carrier lifetime as a function of the dislocation density for n-GaAs. The experimental results are from Yamaguchi et al.[217] They used photoluminescence decay to determine the minority carrier lifetime and TEM to measure the dislocation density. The theoretical curve was calculated assuming $\tau_0 = 2 \times 10^{-8}$ s and $\mu_p = 250\,\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$. It can be seen that the dislocations have a significant effect on the minority carrier lifetime for $D > 10^5\,\text{cm}^{-2}$. In p-GaAs, the minority carriers (electrons) typically have 10 times higher

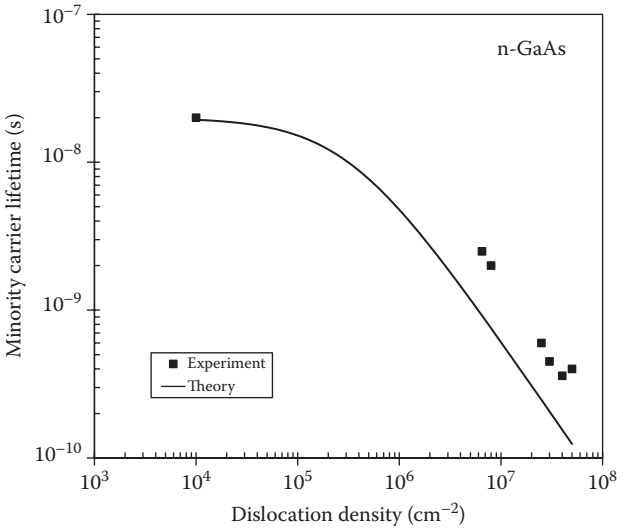


FIGURE 2.33
Minority carrier lifetime as a function of the threading dislocation density in n-GaAs. The experimental results are from Yamaguchi et al.[217] They used photoluminescence decay to determine the minority carrier lifetime and TEM to measure the dislocation density. The theoretical curve was calculated assuming $\tau_0 = 2 \times 10^{-8}$ s and $\mu_p = 250\,\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$. (Data from M. Yamaguchi et al., Analysis for high-efficiency GaAs solar cells on Si substrates, in *20th IEEE Photovoltaic Specialists Conference*, 749 [1988].)

mobility, so this effect is more pronounced. Generally, the minority carrier lifetime will be more affected in materials with high minority carrier mobility.

The Yamaguchi and Amano model has been applied to other dislocated semiconductor materials as well. However, in the case of heteroepitaxial GaN on sapphire (0001), it has been found that only dislocations of screw or mixed character give rise to nonradiative recombination.[214] Edge dislocations, which represent the majority of the threading dislocations in GaN on sapphire (0001), do not appear to contribute to the nonradiative recombination. This explains, in part, why nitride LEDs achieve good performance despite their high threading dislocation densities of ~10^{10} cm^{-2}.[162,218]

In the 3C, 4H, and 6H polytypes of SiC, threading dislocations appear to introduce nonradiative recombination centers, as they do in zinc blende crystals. This results in excess reverse leakage and soft reverse breakdown characteristics in Schottky diodes.[219] This conclusion is reinforced by electron beam–induced current measurements on 4H-SiC p-n junction diodes, which show dark spots associated with dislocations.[220] In the same study, however, bright halos seen around the dislocations indicated that impurities had been gettered from the surrounding material. It is possible that the nonradiative recombination at dislocations is partly or entirely extrinsic (associated with gettered impurities).

It has also been found that threading dislocations degrade carrier mobility in high-electron-mobility transistors (HEMTs), also known as modulation-doped field effect transistors (MODFETs). Ismail[221] studied the effect of threading dislocations on the electron mobility in strained modulation-doped SiGe/Si HEMTs grown on Si substrates by ultrahigh vacuum vapor phase epitaxy. The threading dislocation density was varied systematically by changing the grading rate in the SiGe buffer layer. The dislocation densities were determined by TEM and the carrier mobilities were found using van der Pauw[222] measurements in the temperature range of 0.4–300K. Ismail found that the low-temperature electron

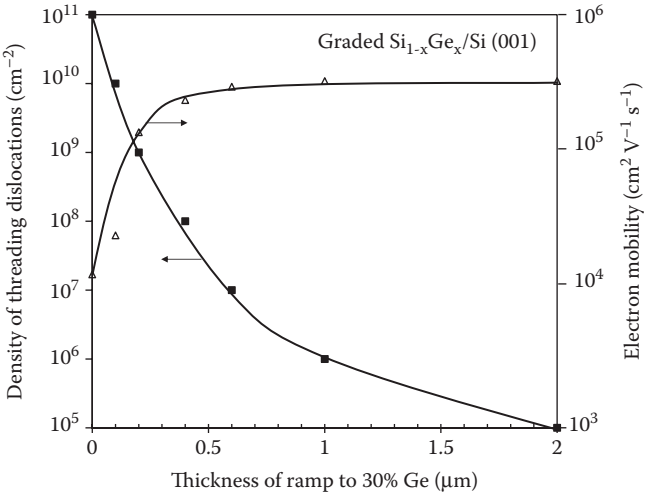


FIGURE 2.34
Effect of the grading rate in the SiGe buffer on the threading dislocation density and the 0.4K electron mobility in a modulation-doped Si/SiGe structure. The structures investigated were grown by UHV-VPE at 500°C–560°C, as follows. On the Si substrate was grown a graded $Si_{1-x}Ge_x$ buffer, with a top composition of $x = 0.3$, followed by a relaxed 1 μm $Si_{0.7}Ge_{0.3}$ buffer, a strained 8–15 nm thick Si channel, and an undoped 15 nm $Si_{1-x}Ge_x$ spacer; an n-type $Si_{1-x}Ge_x$ supply layer; and a 4 nm thick Si cap layer. (Adapted from K. Ismail, *J. Vac. Sci. Technol. B*, 14, 2776 [1996]. With permission. Copyright 1996, American Vacuum Society.)

mobility was degraded by threading dislocations when its density exceeded $3 \times 10^8\,cm^{-2}$, and the mobility was decreased by about two orders of magnitude with a dislocation density of $1 \times 10^{11}\,cm^{-2}$. These results are shown in Figure 2.34. At room temperature, the electron mobility was decreased by 10% and 50%, respectively, with these two dislocation densities. In the same study, it was also found that misfit dislocations in the graded buffer reduce the electron mobility in the Si channel if it is less than 0.4 μm from the buffer.

2.7 Planar Defects

A number of planar crystal defects are encountered in semiconductor heteroepitaxy. These include stacking faults, twins, antiphase domain boundaries (ADBs), and inversion domain boundaries (IDBs).

2.7.1 Stacking Faults

A perfect crystal can be considered a stack of atomic layers occurring in a particular sequence. As explained in Section 2.2.4, the stacking of the zinc blende structure in the [111] direction can be described as …ABCABC…. A stacking fault can occur with an extra plane of atoms inserted into the stacking sequence, as in …ABCBABC…. This is called an extrinsic stacking fault. Another possible type of stacking fault involves the removal of one plane, as in …ABCBC…, and is called an intrinsic stacking fault. Both types of stacking faults are illustrated in Figure 2.35.

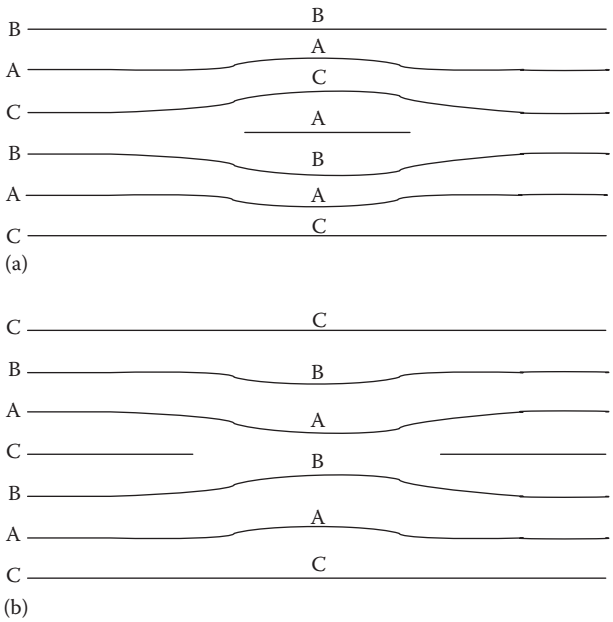


FIGURE 2.35
Schematic illustration of stacking faults in a zinc blende crystal. The stacking direction is [111]. (a) Extrinsic stacking fault. (b) Intrinsic stacking fault.

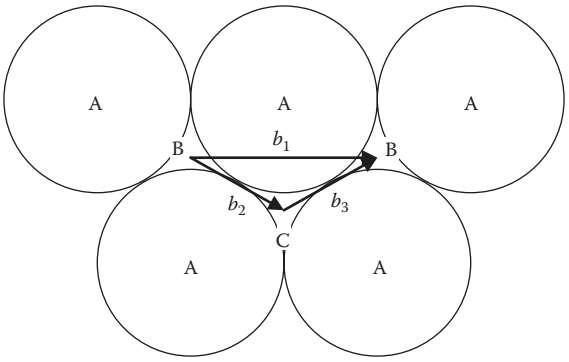


FIGURE 2.36
Schematic drawing of glide on (111) planes by a perfect dislocation (b_1) and by Shockley partial dislocations (b_2 and b_3).

Stacking faults are planar defects that are bounded on either side by partial dislocations. These are called partial dislocations because the Burgers vector is not a lattice translation vector. In other words, the Burgers vector does not start and end on normal lattice sites of the perfect crystal lattice.

Stacking faults may be created during mismatched heteroepitaxial growth,[223] but are also the natural consequence of the dissociation of perfect dislocations into partial dislocations. This occurs naturally during the glide of dislocations, as can be shown with the aid of Figure 2.36. This figure shows the lattice positions on a (111) plane of a zinc blende crystal, labeled A, along with the lattice sites of the underlying and overlying planes, labeled B and C, respectively. The unit of slip (Burgers vector) for a perfect dislocation in the overlying layer is vector b_1. Using the hard sphere model for atoms, this translation takes a sphere in one B position directly to the next B position. However, such a hard sphere will more easily slide first to a C position and then to a B position, along the valleys between the A spheres. These translations are represented by the vectors b_2 and b_3, respectively. Thus, the perfect lattice translation b_1 is naturally split into two simpler translations, $b_2 + b_3$, which are the Burgers vectors associated with two partial dislocations.

In diamond and zinc blende crystals, the perfect 60° dislocation may dissociate into two Shockley partial dislocations by the reaction

$$\frac{a}{2}[011] \rightarrow \frac{a}{6}[112] + \frac{a}{6}[\bar{1}21]. \tag{2.194}$$

The total Burgers vector is conserved as required, and the dissociation is energetically favorable according to Frank's rule. The two partial dislocations lie on the same glide plane as the perfect dislocation; in the example given, this is the $(11\bar{1})$ plane. Between the two partial dislocations, a stacking fault exists on this glide plane.

The equilibrium width of a stacking fault (the separation of the two Shockley partials) may be estimated as follows. Assuming an isotropic crystal, the partial dislocations repel one another with a force per unit length F, given by

$$F = \frac{Gb^2(2-\nu)}{8\pi(1-\nu)d}. \tag{2.195}$$

where G is the shear modulus, b is the length of the Burgers vectors for the partial dislocations, ν is the Poisson ratio, and d is their separation.

The stacking fault that exists between the partial dislocations has an areal energy ξ, which produces an attractive force between the two partials. Equating the two forces, we find the equilibrium width of the stacking fault to be

$$d = \frac{Gb^2}{4\pi\xi} . \tag{2.196}$$

During plastic deformation, two partials making up an extended dislocation will glide simultaneously. As the extended dislocation passes a region of the crystal, the leading partial will create a stacking fault in its wake, after which the trailing partial dislocation will remove this stacking fault.

Often it is adequate in theoretical work to treat extended dislocations as perfect dislocations, and this greatly simplifies the mathematics. This approach is justified if the stacking fault energy is high, causing the dissociated components of the extended dislocation to be tightly bound. Such an assumption is usually good for diamond and zinc blende crystals. In GaAs, for example, the stacking fault energy is 48 mJ/m^2, resulting in an equilibrium stacking fault width of 14.5 nm. For more ionic crystals, the stacking fault energy is less; for example, the value for ZnSe[224] is only 10 mJ/m^2, with a corresponding equilibrium stacking fault width of 70 nm. Often, the observed stacking fault widths are considerably less than the equilibrium values, because the creation of stacking faults requires long-range self-diffusion, which is slow at typical growth temperatures.

2.7.2 Twins

Another type of planar defect resulting from a change in the stacking sequence is the twin. In diamond and zinc blende crystals, twinning occurs almost exclusively on (111) planes. Using the stacking notation of Section 2.2.4, a twin boundary in a diamond or zinc blende crystal may be denoted as …ABCABACBA…. Here the normal crystal and its twin share a single plane of atoms (the "twinning plane" or "composition plane") and there is reflection symmetry about the twinning plane.

Twinning involves a change in long-range order of the crystal; it therefore cannot result from the simple insertion or removal of an atomic plane, as in the case of the stacking fault. Thus, twins cannot be created by the glide of dislocations. Instead, twinning occurs during crystal growth, either bulk growth or heteroepitaxy.

There is a change in crystal orientation at the twinning plane. For a diamond or zinc blende crystal, twinning occurs about a {111} plane. If the original growth plane was (001), then the surface of the twinned crystal is the $(22\bar{1})$ plane.[225] Additional twinning may bring the surface to various $\{22\bar{1}\}$ planes or back to the (001). The (111) twinning plane is always inclined by 54.7° to the (001) surface and may grow out of the crystal in the case of Czochralski-grown bulk crystals.

For (111) growth of zinc blende crystals, the twinned crystal may have either $(\bar{1}\,\bar{1}\,\bar{1})$ or (115) orientation, depending on the orientation of the twinning plane. In the $(\bar{1}\,\bar{1}\,\bar{1})$ case, the twinning plane is the same as the growth plane and will not grow out of the crystal. This is a disadvantage of using the (111) orientation for the growth of bulk crystals.

Twin boundaries are commonly found in heteroepitaxial II-VI crystals grown on (111) substrates, including ZnSe/GaAs (111),[226] CdTe/GaAs (111),[227] and $Hg_xCd_{1-x}Te$/CdTe (111).[228] Twin boundaries are gross defects that degrade device performance, and so (001) is the preferred orientation for heteroepitaxial growth of II-VI materials.

2.7.3 Inversion Domain Boundaries

IDBs, also known as ADBs, are an important consideration in the heteroepitaxy of a polar semiconductor on a nonpolar substrate.[229] Examples include GaN/α-Al_2O_3 (0001), AlN/Si (001), GaAs/Si (001), and InP/Si (001). Due to the lower symmetry of the polar semiconductor, it can grow with one of two (nonequivalent) crystal orientations on the nonpolar substrate. The boundaries between regions having these two orientations are IDBs. Generally, IDBs are expected to introduce states within the energy gap and give rise to nonradiative recombination. They therefore degrade the efficiencies of LEDs and cause excess leakage in p-n junctions. In addition, the charging of IDBs will give rise to scattering of charge carriers and degrade the performance of majority carrier devices such as FETs.

In the case of heteroepitaxial GaN/α-Al_2O_3 (0001), inversion domains have been found in material grown by MOVPE[230] and molecular beam epitaxy (MBE).[231] These manifest as hexagonal domains, 5–20 nm in lateral size, which exist through the entire thickness of the epitaxial layer.[232] The sidewalls of these domains, which are IDBs, are along first-order prism planes of type $\{10\bar{1}0\}$; since these are parallel with the [0001] growth direction, they will not "grow out" of the layer as it is increased in thickness. In the case of MBE-grown GaN on sapphire, convergent beam electron diffraction (CBED) has been used to show that the surfaces of the domains are (0001)N, whereas the surrounding material has (0001)Ga orientation.[233]

For the heteroepitaxial growth of a zinc blende semiconductor on a (001) surface, IDBs that are inclined to the interface may annihilate one another. One such annihilation reaction can occur by the interaction of IDBs on {111} planes,[233] as shown in Figure 2.37. Here, IDB annihilation occurs at the line of intersection of the two IDB planes, which lies along a ⟨110⟩ direction parallel to the interface. Annihilation reactions can also occur between

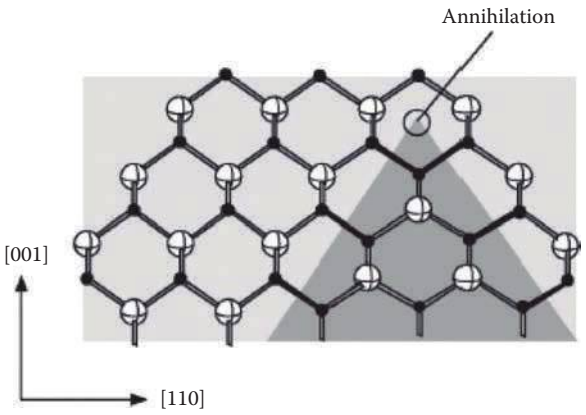


FIGURE 2.37
Annihilation of IDBs on {111} planes in a (001) zinc blende semiconductor. (Reprinted from Y. Ishida et al., *J. Appl. Phys.*, 94, 4676 [2003]. With permission. Copyright 2003, American Institute of Physics.)

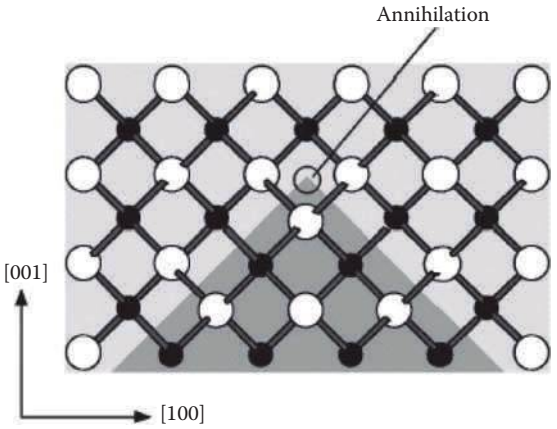


FIGURE 2.38
Annihilation of IDBs on {011} planes in a (001) zinc blende semiconductor. (Reprinted from Y. Ishida et al., *J. Appl. Phys.*, 94, 4676 [2003]. With permission. Copyright 2003, American Institute of Physics.)

IDBs on {011} planes,[234] which can meet along a ⟨010⟩ direction, as shown in Figure 2.38. In either case, the material grown above the annihilation point is free from IDBs. These annihilation mechanisms may be important in 3C-SiC. In zinc blende III-V semiconductors, the IDBs usually exhibit very irregular structures, so that these reaction mechanisms can only occur on a very limited basis.

In zinc blende III-V semiconductors on Si or Ge substrates, inversion domains typically have irregular shapes.[235] The majority of their boundaries do not correspond to low-index crystalline directions. As an example, Figure 2.39 shows IDBs in GaAs on silicon-on-insulator (SOI) (001) grown by MOVPE,[236] revealed by etching[237] in 10:1 HF:HNO_3 and viewed using scanning electron microscopy (SEM). Only small segments of the boundary lines orient along the [110], [010], [120], and occasionally, [130] and [140] directions. The epitaxial film was removed from the SOI using an HF etch. This allowed viewing of the interface side of the GaAs and revealed that the IDBs pass through the entire thickness of the layer. (The inversion domains nucleate at the GaAs/Si interface.)

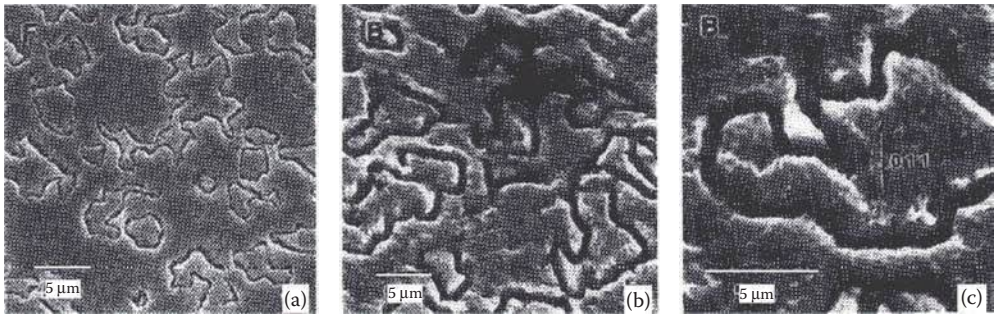


FIGURE 2.39
SEM micrographs of IDBs revealed in a 1.9 μm film of GaAs on SOI. (a) Front side of epitaxial film. (b) Back side of epitaxial film. (c) Higher-magnification micrograph showing [011]-oriented textures. (Reprinted from S. N. G. Chu et al., *J. Appl. Phys.*, 64, 2981 [1988]. With permission. Copyright 1988, American Institute of Physics.)

2.8 Electronic Properties of Semiconductors

The applications of heteroepitaxial materials are ultimately driven by their electronic properties. Some of the most important such properties are tabulated below, including the minimum energy gap at 300K and 0K with its character (direct or indirect), the carrier mobilities and effective masses at 300K, and the relative permittivity (Tables 2.19 and 2.20).

PROBLEMS

1. For Si, calculate the interplanar spacing and atomic density (in atoms cm^{-2}) for each of the following types of planes: (a) {001}, (b) {011}, and (c) {111}.
2. For GaN, calculate the interplanar spacing and atomic density (in atoms cm^{-2}) for each of the following planes: (a) (0001) basal plane, (b) $(10\bar{1}0)$ first-order prism plane, and (c) $(10\bar{1}1)$ first-order pyramidal plane.
3. Determine the compositions of $In_xGa_{1-x}As_yP_{1-y}$, which are lattice matched to GaAs.

TABLE 2.19

Electronic Properties of Cubic Semiconductors

	Energy Gap (eV)		Indirect (I)/Direct (D)	Mobility at 300K		Effective Masses		Permittivity
	300K	0K		μ_n	μ_p	m_n^*/m_0	m_p^*/m_0	$\varepsilon/\varepsilon_0$
C[238]	5.47	5.48	I	1,800	1,200	0.2	0.25	5.7
Si[239]	1.12	1.17	I	1,450	500	$0.98^{l}/0.19^{t}$	$0.16^{lh}/0.49^{hh}$	11.9
Ge[239]	0.66	0.74	I	3,900	1,900	$1.64^{l}/0.082^{t}$	$0.04^{lh}/0.28^{hh}$	16.0
AlAs[239]	2.36	2.23	I	180		0.11	0.22	10.1
AlP[239]	2.42	2.51	I	60	450	0.212	0.145	9.8
AlSb[239]	1.58	1.68	I	200	450	0.12	0.98	14.4
BN[239]	6.4	—	I	200	500	0.26	0.36	7.1
BP[239]	2.0	—	I	40	500	0.67	0.042	11
GaAs[239]	1.42	1.52	D	8,000	400	0.063	$0.076^{lh}/0.5^{hh}$	12.9
GaP[239]	2.26	2.34	I	110	75	0.82	0.60	11.1
GaSb[239]	0.72	0.81	D	5,000	850	0.042	0.40	15.7
InAs[239]	0.36	0.42	D	33,000	460	0.023	0.40	15.1
InP[239]	1.35	1.42	D	4,600	150	0.077	0.64	12.6
InSb[239]	0.17	0.23	D	80,000	1,250	0.0145	0.40	16.8
CdS[239]	2.5	—	D	—	—	0.14	0.51	5.4
CdSe[239]	1.70	1.85	D	800	—	0.13	0.45	10.0
CdTe[239]	1.56	—	D	1,050	100	—	—	10.2
HgSe[239]	−0.30	—	D	20,000		—	—	—
HgTe[240]	−0.15	—	D	25,000	—	—	—	—
ZnS[239]	3.66	3.84	D	600	—	0.39	0.23	8.4
ZnSe[240]	2.58	—	D	—	—	—	—	8.35
ZnTe[240]	2.26	—	D	340	100	—	—	12.67

Note: m_n^*/m_0 is the normalized conductivity effective mass for electrons and m_p^*/m_0 is the normalized conductivity effective mass for holes. t = transverse; l = longitudinal; lh = light hole; hh = heavy hole. The (-) sign in front of the value for the energy gap implies that the compound acts as a semimetal and therefore there exists an overlapping of the valence and conduction band.

TABLE 2.20

Electronic Properties of Hexagonal Semiconductors

	Energy Gap (eV)		Indirect (I)/ Direct (D)	Mobility at 300K		Effective Masses		Permittivity
	300K	0K		μ_n	μ_p	m_n^*/m_0	m_p^*/m_0	$\varepsilon/\varepsilon_0$
2H-SiC	3.33[240]	—	I	—	—	—	—	9.66
4H-SiC	3.263[241]	—	I	450[241]	115[242]	—	—	9.66
6H-SiC	3.023[241]	—	I	500[242]	101[242]	—	—	9.66[242]
AlN	6.03[243]	6.13[244]	D	—	—	0.27[244]	—	8.5[245]
BN	6.1[246]	—	I	—	—	—	—	—
GaN[239]	3.44	3.50	D	400	10	0.27	0.8	10.4
InN	0.641[247]	0.690[248]	D	—	—	0.11[249]	—	—
CdS[239]	2.49	2.58[250]	D	350	40	0.20	0.7	9.1
ZnS[239]	3.66	3.84	D	600	—	0.39	0.23	8.4

Note: m_n^*/m_0 is the normalized conductivity effective mass for electrons and m_p^*/m_0 is the normalized conductivity effective mass for holes. t = transvese; I = longitudinal; lh = light hole; hh = heavy hole.

4. Considering the thermal expansion, find the relaxed lattice constant for GaAs at the following temperatures: (a) 450°C, (b) 550°C, and (c) 650°C.
5. Find the lattice mismatch strain for Ge grown on Si at (a) room temperature and (b) 600°C.
6. Calculate the strain energy per unit area for a pseudomorphic layer of $Si_{0.95}Ge_{0.05}$/Si (001) that is 100 nm thick.
7. Find the line energy per unit length for an edge dislocation in GaAs.
8. Repeat Problem 7 for a screw dislocation in GaAs.
9. Estimate the minority carrier lifetime in n-InP on Si (001) with a dislocation density of 10^8 cm^{-2}, assuming a reasonable value for the minority carrier mobility. Repeat for p-InP/Si (001) having the same dislocation density.

References

1. C. Kittel, *Introduction to Solid State Physics*, 8th ed. (New York: John Wiley & Sons, 2004).
2. R. M. Martin, Relation between elastic tensors of wurtzite and zinc-blende structure materials, *Phys. Rev. B*, 6, 4546 (1972).
3. G. R. Fisher and P. Barnes, Towards a unified view of polytypism in silicon carbide, *Phil. Mag. B*, 61, 217 (1990).
4. K. Jarrendahl and R. B. Davis, in *Semiconductors and Semimetals*, Vol. 52: *SiC Materials and Devices* (New York: Academic Press, 1998), p. 3.
5. J. A. Powell, P. Pirouz, and W. J. Choyke, in *Semiconductor Interfaces, Microstructures and Devices: Properties and Applications*, ed. Z. C. Feng (Bristol: Institute of Physics Publishing, 1993), p. 257.
6. L. S. Ramsdell, Studies on silicon carbide, *Am. Mineral.*, 32, 64 (1947).
7. S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matshusita, H. Kiyoku, and Y. Sugimoto, Ridge-geometry InGaN multi-quantum-well-structure laser diodes, *Appl. Phys. Lett.*, 69, 1477 (1996).
8. S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matshusita, H. Kiyoku, and Y. Sugimoto, InGaN-based multi-quantum-well-structure laser diodes, *Jpn. J. Appl. Phys.*, 35, L74 (1996).

9. J. Bai, T. Wang, H. D. Li, N. Jiang, and S. Sakai, (0001) oriented GaN epilayer grown on $(11\bar{2}0)$ sapphire by MOCVD, *J. Cryst. Growth*, 231, 41 (2001).
10. J. S. Speck and S. F. Chichibu, Nonpolar and semipolar group III nitride-based materials, *MRS Bull.*, 34, 304 (2009).
11. W. von Muench and I. Pfaffender, Thermal oxidation and electrolytic etching of silicon carbide, *J. Electrochem. Soc.*, 122, 642 (1975).
12. A. Suzuki, H. Ashida, N. Furui, K. Maneno, and H. Matsunami, Thermal oxidation of SiC and electrical properties of Al-SiO_2-SiC MOS structure, *Jpn. J. Appl. Phys.*, 21, 579 (1982).
13. H. S. Hong, J. T. Glass, and R. F. Davis, Growth rate, surface morphology, and defect microstructures of β-SiC films chemically vapor deposited on 6H-SiC substrates, *J. Mater. Res.*, 4, 204 (1989).
14. H. Matsunami and T. Kimoto, Step-controlled epitaxial growth of SiC: High quality homoepitaxy, *Mater. Sci. Eng. Rep.*, 20, 125 (1997).
15. T. Kimoto and H. Matsunami, in *Silicon Carbide—Materials, Processing, and Devices*, ed. Z. C. Feng and J. H. Zhao (New York: Taylor and Francis, 2004), p. 1.
16. W. Kaiser and W. L. Bond, Nitrogen, a major impurity in common type I diamond, *Phys. Rev.*, 115, 857 (1959).
17. Y. S. Touloukian, R. K. Kirby, R. E. Taylor, and P. D. Desai, eds., *Thermophysical Properties of Matter*, Vol. 12: *Thermal Expansion, Metallic Elements and Alloys* (New York: Plenum, 1975).
18. E. R. Cohen and B. N. Taylor, The 1986 adjustment of the fundamental physical constants, report of the Committee on Data for Science and Technology of the International Council of Scientific Unions (CODATA) Task Group on Fundamental Constants, CODATA Bulletin 63 (Elmsford, NY: Pergamon, 1986).
19. J. Donahue, *The Structure of the Elements* (New York: John Wiley & Sons, 1974).
20. J. Thewlis and A. R. Davey, Thermal expansion of grey tin, *Nature*, 174, 1011 (1959).
21. A. Taylor and R. M. Jones, in *Silicon Carbide—A High Temperature Semiconductor*, ed. J. R. O'Connor and J. Smiltens (Oxford: Pergamon Press, 1960), p. 147.
22. Y. M. Tairov and V. F. Tsvetkov, Progress in controlling the growth of polytypic crystals, in *Crystal Growth and Characterization of Polytype Structures*, ed. P. Krishna (Oxford: Pergamon Press, 1983), p. 111.
23. R. C. Weast, ed., *Handbook of Chemistry and Physics*, 56th ed. (Cleveland, OH: CRC Press, 1975).
24. Y. S. Touloukian, R. K. Kirby, R. E. Taylor, and T. Y. R. Lee, eds., *Thermophysical Properties of Matter*, Vol. 13: *Thermal Expansion, Nonmetallic Solids* (New York: Plenum, 1977).
25. R. C. Weast, ed., *CRC Handbook of Chemistry and Physics*, 58th ed. (West Palm Beach, FL: CRC Press, 1977).
26. C. C. Wang, M. Zaheeruddin, and L. H. Spinar, Preparation and properties of aluminum phosphide, *J. Inorg. Nucl. Chem.*, 25, 326 (1961).
27. W. B. Pearson, *A Handbook of Lattice Spacings and Structure of Metals and Alloys* (Oxford: Pergamon Press, 1967).
28. G. Giesecke and H. Pfister, Prazionbestimmung der Gitterkonstanten von $A_{III}B_V$—Verbingungen, *Acta Cryst.*, 11, 369 (1958).
29. H. C. Casey Jr. and F. A. Trumbore, Single crystal electroluminescent materials, *Mater. Sci. Eng.*, 6, 69 (1970).
30. E. M. H. Driscoll, A. F. W. Willoughby, J. B. Mullin, and B. W. Straughan, in *Gallium Arsenide and Related Compounds* (London: Institute of Physics). Volumes in Institute of Physics Conference Series, 24, 275 (1975).
31. M. E. Straumanis and C. D. Kim, Lattice parameters, thermal expansion coefficients, phase width, and perfection of the structure of GaSb and InSb, *J. Appl. Phys.*, 36, 3822 (1965).
32. I. Kudman and R. J. Paff, Thermal expansion of $In_xGa_{1-x}As$ alloys, *J. Appl. Phys.*, 43, 3760 (1972).
33. J. V. Ozolin'sh, G. K. Averkieva, A. F. Ilvin'sh, and N. A. Goryunova, An X-ray study of some A^3B^5 compounds which display deviations from stoichiometry, *Sov. Phys. Cryst.* (Engl. transl.), 7, 691 (1963).

34. M. Neuberger, *Handbook of Electronic Materials*, Vol. 2: *III-V Semiconductor Compounds* (New York: Plenum, 1971).
35. J. C. Jamieson and H. H. Demarest, A note on the compression of cubic ZnS, *J. Phys. Chem. Solids*, 41, 963 (1980).
36. W. L. Roth, in *Physics and Chemistry of II-VI Compounds*, ed. M. Aven and J. S. Prener (Amsterdam: North-Holland Publishing, 1975).
37. B. Greenberg, private communication.
38. R. W. G. Wyckoff, *Crystal Structure* (New York: John Wiley & Sons, 1963), p. 108.
39. J. C. Woolley and B. Ray, Solid solution in $A^{II}B^{VI}$ tellurides, *J. Phys. Chem. Solids*, 13, 151 (1960).
40. R. T. Delves and B. Lewis, Zinc blende type HgTe-MnTe solid solutions—I, *J. Phys. Chem. Solids*, 24, 549 (1963).
41. V. Grossman, M. Behringer, H. Heinke, and D. Hommel, Temperature-dependent measurements on ZnSe heterostructures by high-resolution x-ray diffraction, *J. Cryst. Growth*, 184/185, 100 (1998).
42. M. Leszczynski and J. F. Walker, Thermal expansion of gallium arsenide layers grown by molecular beam epitaxy at low temperatures, *Appl. Phys. Lett.*, 62, 1484 (1993).
43. R. Feder and I. Light, Precision thermal expansion measurements of semi-insulating GaAs, *J. Appl. Phys.*, 39, 4870 (1968).
44. E. D. Pierron, D. L. Parker, and J. B. McNeely, Coefficient of expansion of GaAs, GaP, and Ga(As,P) compounds from –62°C to 200°C, *J. Appl. Phys.*, 38, 4669 (1967).
45. S. I. Novikova, Investigation of thermal expansion of GaAs and ZnSe, *Sov. Phys. Sol. State*, 3, 129 (1961).
46. C. Giannini, L. Tapfer, T. Peluso, N. Lovergine, and L. Vasanelli, Residual and thermal strain of ZnS epitaxial layers grown on (100)-GaAs by vapour phase epitaxy, *J. Phys. D*, 28, A125 (1995).
47. J. S. Browder and S. S. Ballard, Low temperature thermal expansion measurements on optical materials, *Appl. Optics*, 8, 793 (1969).
48. R. B. Roberts, G. K. White, and T. M. Sabine, Thermal expansion of zinc sulfide: 300–1300K, *Aust. J. Phys.*, 34, 701 (1981).
49. S. S. Ballard, S. E. Brown, and J. S. Browder, Measurements of the thermal expansion of six optical materials, from room temperature to 250°C, *Appl. Optics*, 17, 1152 (1978).
50. H. P. Singh and B. Dayal, X-ray determination of the thermal expansion of zinc selenide, *Phys. Status Solidi B*, 23, K93 (1967).
51. L. S. Ladd, R. Clark, and D. Roy, U.S. Air Force Technical Report AL-TDR-64-9 (1964).
52. H. P. Maruska and J. J. Tietjen, The preparation and properties of vapor-deposited single-crystalline GaN, *Appl. Phys. Lett.*, 15, 327 (1969).
53. M. Leszczynski, T. Suski, H. Teisseyre, P. Perlin, I. Grzegory, J. Jun, S. Porowski, and T. D. Moustakas, Thermal expansion of gallium nitride, *J. Appl. Phys.*, 76, 4909 (1994).
54. Y. Oshima, T. Suzuki, T. Eri, Y. Kawaguchi, K. Watanabe, M. Shibata, and T. Mishima, Thermal and optical properties of bulk GaN crystals fabricated through hydride vapor phase epitaxy with void-assisted separation, *J. Appl. Phys.*, 98, 103509 (2005).
55. Y. V. Shvyd'ko, M. Lucht, E. Gerdau, M. Lerche, E. E. Alp, W. Sturhahn, J. Sutter, and T. S. Toellner, Measuring wavelengths and lattice constants with the Mössbauer wavelength standard, *J. Synchrotron Rad.*, 9, 17 (2002).
56. Z. Li and R. C. Bradt, Thermal expansion and thermal expansion anisotropy of SiC polytypes, *J. Am. Cer. Soc.*, 70, 445 (1987).
57. Z. Li and R. C. Bradt, Thermal expansion of the hexagonal (4H) polytype of SiC, *J. Appl. Phys.*, 60, 612 (1986).
58. C. Kim, I. K. Robinson, J. Myoung, K. Shim, M.-C. Yoo, and K. Kim, Critical thickness of GaN thin films on sapphire (0001), *Appl. Phys. Lett.*, 69, 2358 (1996).
59. H. Iwanaga, A. Kunishige, and S. Takeuchi, Anisotropic thermal expansion in wurtzite-type crystals, *J. Mater. Sci.*, 35, 2451 (2000).

60. K. Shimada, T. Sota, and K. Suzuki, First-principles study on electronic and elastic properties of BN, AlN, and GaN, *J. Appl. Phys.*, 84, 4951 (1998).
61. S. Poroski, Bulk and homoepitaxial GaN-growth and characterisation, *J. Cryst. Growth*, 189–190, 153 (1998).
62. M. Leszczynski, H. Teisseyre, T. Suski, I. Grzegory, M. Bockowski, J. Jun, K. Pakula, J. M. Baranowski, C. T. Foxon, and T. S. Cheng, Lattice parameters of gallium nitride, *Appl. Phys. Lett.*, 69, 73 (1996).
63. A. U. Sheleg and V. A. Savastenko, Determination of elastic constants of hexagonal crystals from measured values of dynamic atomic displacements, *Izv. Akad. Nauk. BSSR Ser. Fiz. Mat. Nauk.*, 3, 126 (1976).
64. K. Wang and R. R. Reeber, Thermal expansion and elastic properties of InN, *Appl. Phys. Lett.*, 79, 1602 (2001).
65. S. Q. Wang, First-principles study of the anisotropic thermal expansion of wurtzite ZnS, *Appl. Phys. Lett.*, 88, 061902 (2006).
66. V. S. Harutyunyan, A. P. Aivazyan, E. R. Weber, Y. Kim, Y. Park, and S. G. Subramanya, High-resolution x-ray diffraction strain-stress analysis of GaN/sapphire heterostructures, *J. Phys. D Appl. Phys.*, 34, A35 (2001).
67. S. Kalisetty, M. Gokhale, K. Bao, J. E. Ayers, and F. C. Jain, The influence of impurities on the dislocation behavior in heteroepitaxial ZnSe on GaAs, *Appl. Phys. Lett.*, 68, 1693 (1996).
68. S. K. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, 2nd ed. (New York: John Wiley & Sons, 1994), p. 61.
69. S. I. Cho, K. Chang, and M. S. Kwon, Strain analysis of a GaN epilayer grown on a c-plane sapphire substrate with different growth times, *J. Mater. Sci.*, 42, 3569 (2007).
70. L. Vegard, Die constitution der mischkristalle und die raumfullung der atome, *Z. Phys.*, 5, 17 (1921).
71. H. Okuyama, T. Miyajima, Y. Morinaga, F. Hiei, M. Ozawa, and K. Akimoto, ZnSe/ZnMgSSe blue laser diode, *Electron. Lett.*, 28, 1798 (1992).
72. R. L. Moon, G. A. Antypas, and L. W. James, Bandgap and lattice constant of GaInAsP as a function of alloy composition, *J. Electron. Mater.*, 3, 635 (1974).
73. A. N. Baranov, A. N. Imenkov, V. V. Sherstnev, and Yu. P. Yakovlev, 2.7–3.9 μmInAsSb(P)/InAsSbP low diode threshold lasers, *Appl. Phys. Lett.*, 64, 2480 (1994).
74. M. P. C. M. Krijn, Heterojunction band offsets and effective masses in III-V quaternary alloys, *Semicond. Sci. Technol.*, 6, 27 (1991).
75. J. Weiser, J. L. Hoyt, and J. F. Gibbons, Electron mobility enhancement in strained-Si n-type metal-oxide-semiconductor field-effect transistors, *Electron. Dev. Lett.*, 15, 100 (1994).
76. M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie, and A. Lochtefeld, Strained Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors, *J. Appl. Phys.*, 97, 011101 (2005).
77. H. Huntington, Ultrasonic measurements on single crystals, *Phys. Rev.*, 72, 321 (1947).
78. N. Bouarissa, Compositional dependence of the elastic constants and the Poisson ratio of $Ga_xIn_{1-x}Sb$, *Mater. Sci. Eng. B*, 100, 280 (2003).
79. A. E. Merad, H. Aourag, B. Khelifa, C. Mathieu, and G. Merad, Predictions of the bonding properties in $Cd_{1-x}Zn_xTe$, *Superlat. Microstruct.*, 30, 241 (2001).
80. P. C. Kelires, Microstructural and elastic properties of silicon-germanium-carbon alloys, *Appl. Surf. Sci.*, 102, 12 (1996).
81. M. Reason, X. Weng, W. Ye, D. Dettling, S. Hanson, G. Obeidi, and R. S. Goldman, Stress evolution in GaAsN alloy films, *J. Appl. Phys.*, 97, 103523 (2005).
82. M. H. Grimsditch and A. K. Ramdas, Brillouin scattering in diamond, *Phys. Rev. B*, 11, 3139 (1975).
83. S. P. Nikanorov, Yu. A. Burenkov, and A. V. Stepanov, Elastic properties of Si, *Sov. Phys. Solid State* (Engl. transl.), 13, 2516 (1972).
84. Yu. A. Burenkov, S. P. Nikanorov, and A. V. Stepanov, Elastic properties of germanium, *Sov. Phys. Sol. State* (Engl. transl.), 12, 1940 (1971).

85. D. L. Price, J. M. Rowe, and R. M. Nicklow, Lattice dynamics of grey tin and indium antimonide, *Phys. Rev. B*, 3, 1268 (1971).
86. K. B. Tolpygo, Optical, elastic and piezoelectric properties of ionic and valence crystals with the ZnS type lattice, *Sov. Phys. Sol. State*, 2, 2367 (1961).
87. S. Yu. Davydov and A. V. Solomonov, Elastic properties of gallium and aluminum nitrides, *Tech. Phys. Lett.*, 25, 601 (1999). (Transl. from *Pis'ma Zh. Tekh. Fiz.*, 25, 23 [1999].)
88. J. D. Wiley, *Semiconductors and Semimetals*, Vol. 10, ed. R. K. Willardson and A. C. Beer (New York: Academic Press, 1975).
89. D. I. Bolef and M. Menes, Elastic constants of single-crystal aluminum antimonide, *J. Appl. Phys.*, 31, 1426 (1960).
90. R. Weil, Correction to the elastic constants of AlSb, *J. Appl. Phys.*, 43, 4271 (1972).
91. Y. K. Yogurtcu, A. J. Miller, and G. A. Saunders, Pressure dependence of elastic behaviour and force constants of GaP, *J. Phys. Chem. Sol.*, 42, 49 (1981).
92. R. I. Cottam and G. A. Saunders, The elastic constants of GaAs from 2K to 320K, *J. Phys.*, C6, 2105 (1973).
93. G. Simmons and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook* (Cambridge, MA: MIT Press, 1971).
94. F. S. Hinkernell and W. R. Gayton, Elastic constants of single-crystal indium phosphide, *J. Appl. Phys.*, 37, 462 (1966).
95. D. Gerlich, Elastic constants of single-crystal indium phosphide, *J. Appl. Phys.*, 34, 2915 (1963).
96. I. O. Baskin and G. I. Pereseda, Influence of hydrostatic pressures up to 17 kbar on the elastic moduli of InSb, *Sov. Phys. Sol. State* (Engl. transl.) 16, 2058 (1975).
97. I. M. Sil'vestrova, I. B. Kobyakov, and A. A. Shternberg, Elastic and piezoelectric properties of single crystals of sphalerite (ZnS) grown by the hydrothermal method, *Sov. Phys. Tech. Phys.* (Engl. transl.), 25, 1433 (1980).
98. C. G. Hodgins and J. C. Irwin, Elastic constants of ZnSe, *Phys. Status Solidi A*, 28, 647 (1975).
99. D. Berlincourt, H. Jafe, and L. R. Shiozawa, Electroelastic properties of the sulfides, selenides, and tellurides of zinc, *Phys. Rev.*, 129, 1009 (1963).
100. Yu. Yekilov and A. P. Rusakov, Elastic constants and lattice dynamics of some $A^{II}B^{VI}$ compounds, *Sov. Phys. Sol. State* (Engl. transl.), 13, 956 (1971).
101. K. Kumazaki, Elastic properties and iconicity of zero-gap semiconductors, *Phys. Status Solidi A*, 33, 615 (1976).
102. R. I. Cottam and G. A. Saunders, The elastic behavior of mercury telluride, *J. Phys. Chem. Solids*, 36, 1987 (1975).
103. W. G. Cady, *Piezoelectricity* (New York: McGraw-Hill, 1946).
104. R. Bechman, *Proc. IRE*, 36, 764 (1958).
105. G. L. Harris, Miscellaneous properties of SiC, in *Properties of Silicon Carbide*, ed. G. L. Harris (London: INSPEC, Institution of Electrical Engineers, 1995), p. 9.
106. K. Kamitani, M. Grimsditch, J. C. Nipko, C.-K. Loong, M. Okada, and Kimura, The elastic constants of silicon carbide: A brillouin-scattering study of 4H and 6H SiC single crystals, *J. Appl. Phys.*, 82, 3152 (1997).
107. A. Polian, M. Grimsditch, and I. Grzegory, Elastic constants of gallium nitride, *J. Appl. Phys.*, 79, 3343 (1996).
108. C. Deger, E. Born, H. Angerer, O. Ambacher, M. Stutzmann, J. Hornsteiner, E. Riha, and G. Fischerauer, Sound velocity of $Al_xGa_{1-x}N$ thin films obtained by surface acoustic-wave measurements, *Appl. Phys. Lett.*, 72, 2400 (1998).
109. T. Deguchi, D. Ichiryu, K. Toshikawa, K. Segiguchi, T. Sota, R. Matsuo, T. Azuhata, M. Yamaguchi, T. Yagi, S. Chichibu, and S. Nakamura, Structural and vibrational properties of GaN, *J. Appl. Phys.*, 86, 1860 (1999).
110. V. Yu. Davydov, Yu. E. Kitaev, I. N. Goncharuk, A. N. Smirnov, J. Graul, O. Semchinova, D. Uffmann, M. B. Smirnov, A. P. Mirgorodsky, and R. A. Evarestov, Phonon dispersion and Raman scattering in hexagonal GaN and AlN, *Phys. Rev. B*, 58, 12899 (1998).

111. V. A. Savastenko and A. U. Sheleg, Study of the elastic properties of gallium nitride, *Phys. Status Solidi A*, 48, K135 (1978).
112. L. E. McNeil, M. Grimsditch, and R. H. French, Vibrational spectroscopy of aluminum nitride, *J. Am. Ceram. Soc.*, 76, 1132 (1993).
113. K. Tsubouchi, K. Sugai, and N. Mikoshiba, AlN material constants evaluation and SAW properties on AlN/Al_2O_3 and AlN/Si, in *1981 Ultrasonics Symposium Proceedings*, ed. B. R. McAvoy (New York: IEEE, 1981), p. 375.
114. K. Kim, W. R. L. Lambrecht, and B. Segall, Elastic constants and related properties of tetrahedrally bonded BN, AlN, GaN, and InN, *Phys. Rev. B*, 53, 16310 (1996).
115. A. F. Wright, Elastic properties of zinc-blende and wurtzite AlN, GaN, and InN, *J. Appl. Phys.*, 82, 2833 (1997).
116. A. A. Marmalyuk, R. K. Akchurin, and V. A. Gorbylev, Evaluation of elastic constants of AlN, GaN, and InN, *Inorg. Mater.*, 34, 691 (1998). (Transl. of *Neorganicheskie Materialy*, 34, 833 [1998]).
117. J. A. Chisholm, D. W. Lewis, and P. D. Bristowe, Classical simulations of the properties of group-III nitrides, *J. Phys. Cond. Mater.*, 11, L235 (1999).
118. J.-M. Wagner and F. Bechstedt, Properties of strained wurtzite GaN and AlN: Ab initio studies, *Phys. Rev. B*, 66, 115202 (2002).
119. I. Yonenaga, T. Shima, and M. H. F. Sluiter, Nan-indentation hardness and elastic moduli of bulk single-crystal AlN, *Jpn. J. Appl. Phys.*, 41, Pt. 1, 4620 (2002).
120. V. Y. Davydov, N. S. Averkiev, I. N. Goncharuk, D. K. Nelson, I. P. Nikitna, A. S. Polkovnikov, A. N. Smirnov, M. A. Jacobson, and Q. W. Semchinova, Raman and photoluminescence studies of biaxial strain in GaN epitaxial layers grown on 6H-SiC, *J. Appl. Phys.*, 82, 5097 (1997).
121. B. Gil, *Low-Dimensional Nitride Semiconductors*, 1st ed. (Oxford: Oxford University Press, 2002), p. 26.
122. V. Yu. Davydov, N. S. Averkiev, I. N. Goncharuk, D. K. Nelson, I. P. Nikitina, A. S. Polkovnikov, A. N. Smirnov, and M. A. Jacobson, Raman and photoluminescence studies of biaxial strain in GaN epitaxial layers grown on 6H-SiC, *J. Appl. Phys.*, 82, 5097 (1997).
123. J. J. Gilman, Direct measurement of the surface energies of crystals, *J. Appl. Phys.*, 31, 2208 (1960).
124. D. Hull, P. Beardmore, and A. P. Valentine, Crack propagation in single crystals of tungsten, *Phil. Mag.*, 12, 1021 (1965).
125. A. Kelly, *Strong Solids* (Oxford: Clarendon Press, 1966), p. 45.
126. H. P. Bonzel, Equilibrium crystal shapes: Towards absolute energies, *Prog. Surf. Sci.*, 67, 45 (2001).
127. J. J. Metois and P. Muller, Absolute surface energy determination, *Surf. Sci.*, 548, 13 (2004).
128. R. J. Jaccodine, Surface energy of germanium and silicon, *J. Electrochem. Soc.*, 110, 524 (1963).
129. C. Messmer and J. C. Bilello, The surface energy of Si, GaAs, and GaP, *J. Appl. Phys.*, 52, 4823 (1981).
130. J. W. Cahn and R. E. Hanneman, (111) surface tensions of III-V compounds and their relationship to spontaneous bending of thin crystals, *Surf. Sci.*, 1, 387 (1964).
131. D. Holec and P. H. Mayrhofer, Surface energies of AlN allotropes from first principles, *Scripta Materialia*, 67, 760 (2012).
132. J. E. Northrup and L. T. Romano, Surface energetics, pit formation, and chemical ordering in InGaN alloys, *Appl. Phys. Lett.*, 74, 16 (1999).
133. C. Tang, M. J. S. Spencer, and A. Barnard, Activity of ZnO polar surfaces: An insight from surface energies, *Phys. Chem. Chem. Phys.*, 16, 22139 (2014).
134. J. W. Matthews, Defects associated with the accommodation of misfit between crystals, *J. Vac. Sci. Technol.*, 12, 126 (1975).
135. University of Cambridge, Cambridge, UK, DoITPoMS web site. www.msn.cam.ac.uk.
136. F. A. Ponce, D. Cherns, W. T. Young, and J. W. Steeds, Characterization of dislocations in GaN by transmission electron diffraction and microscopy techniques, *Appl. Phys. Lett.*, 69, 770 (1996).
137. T. Metzger, R. Hopler, E. Born, O. Ambacher, M. Stutzmann, R. Stommer, M. Schuster, H. Gobel, S. Christiansen, M. Albrecht, and H. P. Strunk, Defect structure of epitaxial GaN films determined by transmission electron microscopy and triple-axis x-ray diffractometry, *Phil. Mag. A*, 77, 1013 (1998).

138. S. V. Ivanov, P. D. Altukhov, T. S. Argunova, A. A. Bakun, A. A. Budza, V. V. Chaldyshev, Yu. A. Kovalenko, et al., Molecular beam epitaxy growth and characterization of thin (<2 μm) GaSb layers on GaAs (100) substrates, *Semicond. Sci. Technol.*, 8, 347 (1993).
139. S. N. G. Chu and S. Nakahara, 1/2<100>{100} dislocation loops in a zinc blende structure, *Appl. Phys. Lett.*, 66, 434 (1990).
140. B. C. de Cooman and C. B. Carter, The accommodation of misfit at {100} heterojunctions in III-V compound semiconductors by gliding dissociated dislocations, *Acta Met.*, 37, 2765 (1989).
141. T. Kamejima, K. Ishida, and J. Matsui, Injection-enhanced dislocation glide under uniaxial stress in GaAs-(GaAl)As double heterostructure laser, *Jpn. J. Appl. Phys.*, 16, 233 (1977).
142. E. Peissker, P. Haasen, and H. Alexander, Anisotropic plastic deformation of indium antimonide, *Phil. Mag.*, 7, 1279 (1962).
143. R. L. Bell and A. F. W. Willoughby, Etch-pit studies of dislocations in indium antimonide, *J. Mater. Sci.*, 1, 219 (1966).
144. T. Ninomiya, *J. Phys. Coll. Chem.*, C640, C6 (1979).
145. J. Petruzzello and M. R. Leys, Effect of the sign of misfit strain on the dislocation structure at interfaces of $GaAs_xP_{1-x}$ films, *Appl. Phys. Lett.*, 53, 2414 (1988).
146. M. Yokozeki, H. Yonezu, T. Tsuji, and N. Ohshima, Passivation of misfit dislocations by atomic hydrogen irradiation in lattice mismatched heteroepiaxy, *J. Cryst. Growth*, 175/176, 435 (1997).
147. J. L. Weyher and J. van de Ven, Influence of substrate defects on the structure of epitaxial GaAs grown by MOCVD, *J. Cryst. Growth*, 88, 221 (1988).
148. Wacker-Siltronic, Munich.
149. M/A-Com, Lowell, MA.
150. K. Durose and G. J. Russell, Structural defects in CdTe crystals grown by two different vapour phase techniques, *J. Cryst. Growth*, 86, 471 (1988).
151. N. M. Aitken, M. D. G. Potter, D. J. Buckley, J. T. Mullins, J. Carles, D. P. Halliday, K. Durose, B. K. Tanner, and A. W. Brinkman, Characterisation of cadmium telluride bulk crystals grown by a novel "multi-tube" vapour growth technique, *J. Cryst. Growth*, 198/199, 984 (1999).
152. G. A. Rozgonyi, P. M. Petroff, and M. B. Panish, Control of lattice parameters and dislocations in the system $Ga_{1-x}Al_xAs_{1-y}P_y$/GaAs, *J. Cryst. Growth*, 27, 106 (1974).
153. K. R. Breen, P. N. Uppal, and J. S. Ahearn, Interface dislocation structures in $In_xGa_{1-x}As$/GaAs mismatched epitaxy, *J. Vac. Sci. Technol. B*, 7, 758 (1989).
154. M. J. Matragrano, D. G. Ast, J. R. Shealy, and V. Krishnamoorthy, Anisotropic strain relaxation of GaInP epitaxial layers in compression and tension, *J. Appl. Phys.*, 79, 8371 (1996).
155. S. H. Huang, G. Balakrishnan, A. Khoshaklagh, A. Jallipalli, L. R. Dawson, and D. L. Huffaker, Strain relief by periodic misfit arrays for low defect density GaSb on GaAs, *Appl. Phys. Lett.*, 88, 131911 (2006).
156. S. Srinivasan, L. Geng, R. Liu, and F. A. Ponce, Slip systems and misfit dislocations in InGaN epilayers, *Appl. Phys. Lett.*, 83, 5187 (2003).
157. Y. N. Picard, M. E. Twigg, J. D. Caldwell, C. R. Eddy Jr., M. A. Mastro, and R. T. Holm, Resolving the Burgers vector for individual GaN dislocations by electron channeling contrast imaging, *Scripta Mater.*, 61, 773 (2009).
158. S. Nakamura, GaN growth using GaN buffer layer, *Jpn. J. Appl. Phys.*, 30, L1705 (1991).
159. H. Amano, N. Sawaki, I. Akasaki, and Y. Toyoda, Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer, *Appl. Phys. Lett.*, 48, 353 (1986).
160. D. Kapolnek, X. H. Wu, B. Heying, S. Keller, B. P. Keller, U. K. Mishra, S. P. DenBaars, and J. S. Speck, Structural evolution in epitaxial metalorganic chemical vapor deposition grown GaN films on sapphire, *Appl. Phys. Lett.*, 67, 1541 (1995).
161. W. Qian, M. Skowronski, M. De Graef, K. Doverspike, L. B. Rowland, and D. K. Gaskill, Microstructural characterization of α-GaN films grown on sapphire by organometallic vapor phase epitaxy, *Appl. Phys. Lett.*, 66, 1252 (1995).
162. F. R. Chien, X. J. Ning, S. Stemmer, P. Pirouz, M. D. Bremser, and R. F. Davis, Growth defects in GaN films on 6H-SiC substrates, *Appl. Phys. Lett.*, 68, 2678 (1996).

163. I. Y. Knoke, E. Meissner, J. Friedrich, H. P. Strunk, and G. Muller, Reduction of the dislocation density in GaN during low-pressure solution growth, *J. Cryst. Growth*, 310, 3351 (2008).
164. E. Valcheva, T. Paskova, and B. Monemar, Nanopipes and their relationship to the growth mode in thick HVPE-GaN layers, *J. Cryst. Growth*, 255, 19 (2003).
165. W. Qian, G. S. Rohrer, M. Skowronski, K. Doverspike, L. B. Rowland, and D. K. Gaskill, Open-core screw dislocations in GaN epilayers observed by scanning force microscopy and high-resolution transmission electron microscopy, *Appl. Phys. Lett.*, 67, 2284 (1995).
166. F. A. Ponce, J. S. Major Jr., W. E. Plano, and D. F. Welch, Crystalline structure of AlGaN epitaxy on sapphire using AlN buffer layers, *Appl. Phys. Lett.*, 65, 2302 (1994).
167. H. P. D. Schenk, P. Vennegues, O. Tottereau, T. Riemann, and J. Christen, Three-dimensionally nucleated growth of gallium nitride by low-pressure metalorganic vapour phase epitaxy, *J. Cryst. Growth*, 258, 232 (2003).
168. W. J. Choyke and G. Pensl, Physical properties of SiC, *Mater. Res. Soc. Bull.*, 22, 36 (1997).
169. M. Dudley, S. Wang, W. Huang, C. H. Carter Jr., V. F. Tsvetkov, and C. Fazi, White-beam synchrotron topographic studies of defects in 6H-SiC single crystals, *J. Phys. D*, 28, A63 (1995).
170. B. M. Epelbaum and D. Hofmann, On the mechanisms of micropipe and macrodefect transformation in SiC during liquid phase treatment, *J. Cryst. Growth*, 225, 1 (2001).
171. I. Kamata, H. Tsuchida, T. Jikimoto, and K. Izumi, Structural transformation of screw dislocations via thick 4H-SiC epitaxial growth, *Jpn. J. Appl. Phys.*, 39, 6496 (2000).
172. S. Ha, N. T. Nuhfer, G. S. Rohrer, M. De Graef, and M. Skowronski, Origin of the domain structure in hexagonal silicon carbide boules grown by the physical vapor transport method, *J. Cryst. Growth*, 220, 308 (2000).
173. M. Tuominen, R. Yakimova, R. C. Glass, T. Tuomi, and E. Janzen, Crystal imperfections in 4H SiC grown with a seeded Lely method, *J. Cryst. Growth*, 144, 267 (1994).
174. J. Takahashi, N. Ohtani, and M. Kanaya, Structural defects in α-SiC single crystals grown by the modified-Lely method, *J. Cryst. Growth*, 167, 596 (1996).
175. D. Hobgood, M. Brady, W. Brixius, G. Fetchko, R. Glass, D. Henshall, J. Jenny, et al., Status of large diameter SiC crystal growth for electronic and optical applications, *Mater. Sci. Forum*, 338–342, 3 (2000).
176. M. Dudley and X. Huang, Characterization of SiC using synchrotron white beam x-ray topography, *Mater. Sci. Forum*, 338–342, 431 (2000).
177. Cree Materials, Durham, NC.
178. D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. (Oxford: Butterworth Heinemann, 2001), p. 65.
179. F. R. N. Nabarro, *Theory of Crystal Dislocations* (New York: Dover, 1987), p. 201.
180. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers. I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).
181. A. A. Maradudin, Screw dislocations and discrete elastic theory, *J. Phys. Chem. Solids*, 9, 1 (1958).
182. R. H. M. van der Leur, A. J. G. Schellingerhout, F. Tuinstra, and J. E. Mooij, Critical thickness for pseudomorphic growth of Si/Ge alloys and superlattices, *J. Appl. Phys.*, 64, 3043 (1988).
183. F. Kroupa, Circular edge dislocation loop, *Czech. J. Phys. B*, 10, 284 (1960).
184. A. Bere and A. Serra, Atomic structure of dislocation cores in GaN, *Phys. Rev. B*, 65, 205323 (2002).
185. S. M. Lee, A. Belkhir, X. Y. Zhu, Y. H. Lee, Y. G. Hwang, and Th. Frauenheim, Electronic structure of GaN edge dislocations, *Phys. Rev. B*, 61, 16033 (2000).
186. D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. (Oxford: Butterworth Heinemann, 2001), p. 71.
187. J. W. Steeds, *Introduction to Anisotropic Elasticity Theory of Dislocations* (Oxford: Clarendon Press, 1973).
188. D. Holec, P. M. F. J. Costa, M. J. Kappers, and C. J. Humphreys, Critical thickness calculations for InGaN/GaN, *J. Cryst. Growth*, 303, 314 (2007).
189. I. Belabbas, A. Bere, J. Chen, S. Petit, M. A. Belkhir, P. Ruterana, and G. Nouet, Atomistic modeling of the (a+c)-mixed dislocation core in wurtzite GaN, *Phys. Rev. B*, 75, 115201 (2007).

190. L. Lymperakis, Ab-initio based multiscale calculations of extended defects in and on group III-nitrides, PhD thesis, Department Physik der Fakultat fur Naturwissenschaften an der Universitat Paderborn (2005).
191. J. Kioseoglou, Ph. Komninou, and Th. Karakostas, Core models of a-edge threading dislocations in wurtzite III(Al,Ga,In)-nitrides, *Phys. Status Solidi, A* 206, 1931 (2009).
192. F. C. Frank, Capillary equilibria of dislocated crystals, *Acta Cryst.*, 4, 497 (1951).
193. W. M. Vetter and M. Dudley, Micropipes and the closure of axial screw dislocation cores in silicon carbide crystals, *J. Appl. Phys.*, 96, 348 (2004).
194. D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. (Oxford: Butterworth Heinemann, 2001), p. 72.
195. M. Peach and J. S. Koehler, The forces exerted on dislocations and the stress fields produced by them, *Phys. Rev.*, 80, 436 (1950).
196. D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. (Oxford: Butterworth Heinemann, 2001), p. 42.
197. W. G. Johnston and J. J. Gilman, Dislocation velocities, dislocation densities, and plastic flow in lithium fluoride crystals, *J. Appl. Phys.*, 30, 129 (1959).
198. S. K. Choi, M. Mihara, and T. Ninomiya, Dislocation velocities in GaAs, *Jpn. J. Appl. Phys.*, 16, 737 (1977).
199. S. A. Erofeeva and Yu. A. Osip'yan, *Sov. Phys. Sol. State* (Engl. transl.), 15, 538 (1973).
200. J. J. Gilman, Dislocation mobility in crystals, *J. Appl. Phys.*, 36, 3195 (1965).
201. A. R. Chaudhuri, J. R. Patel, and L. G. Rubin, Velocities and densities of dislocations in germanium and other semiconductor crystals, *J. Appl. Phys.*, 33, 2736 (1962).
202. R. Hull and J. C. Bean, Experimental and theoretical analysis of strain relaxation in Ge_xSi_{1-x}/Si(100) heteroepitaxy, *Mater. Res. Symp. Proc.*, 160, 23 (1990).
203. R. Hull, J. C. Bean, and C. Buescher, A phenomenological description of strain relaxation in Ge_xSi_{1-x}/Si(100) heterostructures, *J. Appl. Phys.*, 66, 5837 (1989).
204. A. George and J. Rabier, Dislocations and plasticity in semiconductors. I. Dislocation structures and dynamics, *Rev. Phys. Appl.*, 22, 941 (1987).
205. I. Yonenaga and K. Sumino, Effects of In impurity on the dynamic behavior of dislocations in GaAs, *J. Appl. Phys.*, 62, 1212 (1987).
206. V. Swaminathan and S. M. Copley, Hardening of GaAs by solute-vacancy pairs, *J. Appl. Phys.*, 47, 4405 (1976).
207. B. W. Dodson, Stress dependence of dislocation glide activation energy in single-crystal silicon-germanium alloys up to 2.6 GPa, *Phys. Rev. B*, 38, 12383 (1988).
208. J. P. Hirth and J. Lothe, *Theory of Crystal Dislocations* (New York: McGraw Hill, 1968).
209. R. Hull, J. C. Bean, D. Bahnk, L. J. Peticolas Jr., K. T. Short, and F. C. Unterwald, Interpretation of dislocation propagation velocities in strained Ge_xSi_{1-x}/Si(100) heterostructures by the diffusive kink model, *J. Appl. Phys.*, 70, 2052 (1991).
210. A. Seeger and P. Schiller, The formation and diffusion of kinks as the fundamental process of dislocation movement in internal friction measurements, *Acta Metall.*, 10, 348 (1962).
211. P. W. Hutchinson and P. S. Dobson, Defect structure of degraded GaAlAs-GaAs double heterostructure lasers, *Phil. Mag.*, 32, 745 (1975).
212. D. V. Lang and C. H. Henry, Nonradiative recombination at deep levels in GaAs and GaP by lattice-relaxation multiphonon emission, *Phys. Rev. Lett.*, 35, 1525 (1975).
213. T. Hino, S. Tomiya, T. Miyajima, K. Yanashima, S. Hashimoto, and M. Ikeda, Characterization of threading dislocations in GaN epitaxial layers, *Appl. Phys. Lett.*, 76, 3421 (2000).
214. S. Tomiya, E. Morita, M. Ukita, H. Okuyama, S. Itoh, K. Nakano, and A. Ishibashi, Structural study of defects induced during current injection to II-VI blue light emitter, *Appl. Phys. Lett.*, 66, 1208 (1995).
215. F. R. N. Nabarro, *Theory of Crystal Dislocations* (New York: Dover, 1987), p. 51.
216. M. Yamaguchi and C. Amano, Efficiency calculations of thin-film GaAs solar cells on Si substrates, *J. Appl. Phys.*, 58, 3601 (1985).

217. M. Yamaguchi, C. Amano, Y. Itoh, K. Hane, R. A. Ahrenkiel, and M. M. Al-Jassim, Analysis for high-efficiency GaAs solar cells on Si substrates, in *20th IEEE Photovoltaic Specialists Conference*, 749 (1988).
218. S. D. Lester, F. A. Ponce, M. G. Craford, and D. A. Steigerwald, High dislocation densities in high efficiency GaN-based light-emitting diodes, *Appl. Phys. Lett.*, 66, 1249 (1995).
219. P. G. Neudeck, W. Huang, and M. Dudley, Breakdown degradation associated with elementary screw dislocations in 4H-SiC p+n junction, *Mater. Res. Soc. Symp. Proc.*, 483, 285 (1998).
220. S. Maximenko, S. Soloviev, D. Cherednichenko, and T. Sudarshan, Electron-beam-induced current observed for dislocations in diffused 4H-SiC PN diodes, *Appl. Phys. Lett.*, 84, 1576 (2004).
221. K. Ismail, Effect of dislocations in strained Si/SiGe on electron mobility, *J. Vac. Sci. Technol. B*, 14, 2776 (1996).
222. L. J. van der Pauw, A method for measuring specific resistivity and Hall effect of discs of arbitrary shape, *Phillips Res. Rep.*, 13, 1 (1958).
223. K. Yamane, T. Kawai, Y. Furukawa, H. Okada, and A. Wakahara, Growth of low defect density GaP layers on Si substrates within the critical thickness by optimized shutter sequence and post-growth annealing, *J. Cryst. Growth*, 312, 2179 (2010).
224. H. Hartmann, R. Mach, and B. Selle, in *Current Topics in Material Science*, ed. E. Kaldis, Vol. 9 (Amsterdam: North-Holland, 1982), p. 1.
225. W. R. Runyan, *Silicon Semiconductor Technology* (New York: McGraw-Hill, 1965), p. 98.
226. T. Yao and S. Maekawa, Molecular beam epitaxy of zinc chalcogenides, *J. Cryst. Growth*, 53, 423 (1981).
227. P. D. Brown, J. E. Hails, G. J. Russell, and J. Woods, Defect structure of epitaxial CdTe layers grown in {100} and {111}B GaAs and on {111}B CdTe by metalorganic chemical vapor deposition, *Appl. Phys. Lett.*, 50, 1144 (1987).
228. J. E. Hails, G. J. Russell, A. W. Brinkman, and J. Woods, The effect of CdTe substrate orientation on the MOVPE growth of $Cd_xHg_{1-x}Te$, *J. Cryst. Growth*, 79, 940 (1986).
229. H. Kroemer, Polar-on-nonpolar epitaxy, *J. Cryst. Growth*, 81, 193 (1987).
230. J. L. Rouviere, M. Arlery, A. Bourret, R. Niebuhr, and K. Bachem, Understanding the pyramidal growth of GaN, *Mater. Res. Soc. Symp. Proc.*, 395, 393 (1996).
231. V. Potin, P. Ruterana, M. Benamara, and H. P. Strunk, Inversion domains and pinholes in GaN grown over Si(111), *Appl. Phys. Lett.*, 82, 4471 (2003).
232. P. Ruterana, Convergent beam electron diffraction investigation of inversion domains in GaN, *J. Alloys Compounds*, 401, 199 (2005).
233. M. Kawabe and T. Ueda, Self-annihilation of antiphase boundary in GaAs on Si(100) grown by molecular beam epitaxy, *Jpn. J. Appl. Phys.*, Pt. 2, 26, L944 (1987).
234. Y. Li and L. J. Giling, Growth by atmospheric pressure OMVPE and x-ray analysis of ZnTe epilayers on III-V substrates, *J. Cryst. Growth*, 163, 203 (1996).
235. K. Morizane, Antiphase domain structures in GaP and GaAs epitaxial layers grown on Si and Ge, *J. Cryst. Growth*, 38, 249 (1977).
236. S. N. G. Chu, S. Nakahara, S. J. Pearton, T. Boone, and S. M. Vernon, Antiphase domains in GaAs grown by metalorganic chemical vapor deposition on silicon-on-insulator, *J. Appl. Phys.*, 64, 2981 (1988).
237. P. N. Uppal and H. Kroemer, Molecular beam epitaxial growth of GaAs on Si (211), *J. Appl. Phys.*, 58, 2195 (1985).
238. S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. (Hoboken, NJ: John Wiley & Sons, 2007).
239. L. I. Berger and B. R. Pamplin, Properties of semiconductors, in *CRC Handbook of Physics and Chemistry*, ed. D. R. Lide, 84th ed. (Boca Raton, FL: CRC Press, 2004).
240. S. Yoshida, Band structure of SiC, in *Properties of Silicon Carbide*, ed. G. L. Harris (Exeter, England: Short Run Press Ltd., 1995).
241. G. L. Harris, H. S. Henry, and A. Jackson, Carrier mobilities and concentrations in SiC, in *Properties of Silicon Carbide*, ed. G. L. Harris (Exeter, England: Short Run Press Ltd., 1995).

242. L. Patrick and W. J. Choyke, Static dielectric constant of SiC, *Phys. Rev. B*, 2, 2255 (1970).
243. Q. Guo and A. Yoshida, Temperature dependence of band gap change in InN and AlN, *Jpn. J. Appl. Phys.*, 33, 2453 (1994).
244. M. Suzuki and T. Uenoyama, First-principles calculations of effective-mass parameters of AlN and GaN, *Phys. Rev. B*, 52, 8132 (1995).
245. O. Ueda and S. J. Pearton, *Materials and Reliability Handbook for Semiconductor Optical and Electron Devices* (New York: Springer Science + Business Media, 2013).
246. S. L. Rumyantsev, M. E. Levinshtein, A. D. Jackson, S. N. Mohammmad, G. L. Harris, M. G. Spencer, and M. S. Shur, Boron Nitride (BN), in *Properties of Advanced Semiconductor Materials GaN, AlN, InN, BN, SiC, SiGe*, ed. M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur (New York: John Wiley & Sons, 2001), p. 67.
247. J. Wu, W. Walukiewicz, W. Shan, K. M. Yu, J. W. Ager III, S. X. Li, E. E. Haller, H. Lu, and W. Schaff, Temperature dependence of the fundamental band gap of InN, *J. Appl. Phys.*, 94, 4457 (2003).
248. W. Walukiewicz, J. W. Ager III, K. M. Yu, Z. Liliental-Weber, J. Wu, S. X. Li, R. E. Jones, and J. D. Denlinger, Structure and electronic properties of InN and In-rich group III-nitride alloys, *J. Phys. D Appl. Phys.*, 39, R83 (2006).
249. H. Morkoc, *Handbook of Nitride Semiconductors and Devices*, Vol. 1: *Materials Properties, Physics and Growth* (Weinheim, Germany: Wiley-VCH Verlag, 2008).
250. M. L. Cohen, J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 2nd ed. (Berlin: Springer-Verlag, 1998 and 1999).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

3

Heteroepitaxial Growth

3.1 Introduction

Of the many available epitaxial growth techniques, molecular beam epitaxy (MBE) and metalorganic vapor phase epitaxy (MOVPE) have emerged as general-purpose tools for heteroepitaxial research and commercial production. This is because these methods afford tremendous flexibility and the ability to deposit thin layers and complex multilayered structures with precise control and excellent uniformity. Together, MBE and MOVPE account for virtually all production of compound semiconductor devices today.

MBE is an ultra-high-vacuum (UHV) technique that involves the impingement of atomic or molecular beams onto a heated single-crystal substrate where the epitaxial layers grow. The source beams originate from Knudsen evaporation cells or gas source crackers. These can be turned on and off very abruptly by shutters and valves, respectively, providing atomic layer abruptness. Because MBE takes place in a UHV environment, it is possible to employ a number of *in situ* characterization tools based on electron or ion beams. These provide the crystal grower with immediate feedback, and improved control of the growth process. Another advantage of MBE is flexibility; nearly all semiconductors can be grown, including III-V and II-VI semiconductors; Si, Ge, and $Si_{1-x}Ge_x$ alloys; and SiC and $Si_{1-x-y}Ge_xC$ alloys. However, III-phosphides are difficult to grow by MBE and alloys involving As and P are especially troublesome. Other drawbacks of MBE are the initial high cost and maintenance requirements of the UHV system, as well as limited throughput.

MOVPE is a vapor phase epitaxial (VPE) process that is carried out at atmospheric or reduced (e.g., 0.1 atm) pressure using metalorganic precursors. Often, hydride sources are used in conjunction with the metalorganic chemicals; occasionally, even elemental sources are used. Like MBE, MOVPE provides excellent control over the growth of thin layers and multilayered structures, including quantum well devices and superlattices. However, the lack of UHV conditions precludes the *in situ* use of electron or ion beam characterization tools. Nonetheless, optical *in situ* characterization methods have been utilized to some extent. Another disadvantage of MOVPE is the use of highly toxic source chemicals, especially arsine (AsH_3) and phosphine (PH_3). These flammable, explosive, and highly toxic gas sources are stored at high pressure in large quantities, raising a number of safety concerns in a production environment. In some cases, these hydride sources have been replaced with less toxic liquid sources (such as TBAs) contained in low-pressure bubblers.

VPE processes are also used for the growth of column IV semiconductors, including Si, Ge, and $Si_{1-x}Ge_x$ alloys and SiC and $Si_{1-x-y}Ge_xC$ alloys. These VPE processes utilize similar equipment and share some of the characteristics of MOVPE, but do not involve metalorganic precursors. Instead, hydride and halide sources are used. The use of all hydride sources leads to an irreversible process with abrupt interfaces. On the other hand, any

involvement of halide precursors generally results in a reversible process that is generally less suitable for the growth of multilayered structures.

This chapter provides a brief overview of the important epitaxial processes of VPE and MBE. MOVPE is considered a special case of VPE. The remaining sections of the chapter describe the growth of particular materials, from the viewpoint of mismatched heteroepitaxy.

3.2 Vapor Phase Epitaxy

VPE growth is accomplished by passing gaseous source chemicals over a heated single-crystal substrate, where epitaxial growth occurs. Atmospheric or reduced (~0.1 atm) pressure may be used. In either case, a carrier gas such as hydrogen usually makes up most of the flow (and therefore pressure) in the reactor.

VPE growth is extremely flexible, and allows the growth of nearly every semiconductor material of interest. The availability of ultrapure sources and careful reactor design allow the growth of materials with levels of purity matching all other epitaxial techniques. It has also proved possible to design reactors capable of handling multiple wafers in one run, while maintaining excellent uniformity, both across wafers and from wafer to wafer. This scalability has led to its wide commercial application.

3.2.1 Vapor Phase Epitaxial Mechanisms and Growth Rates

The VPE growth of a crystal involves a series of basic steps: (1) the source chemicals are transported in the vapor phase to the heated substrate; (2) source molecules diffuse to the growing surface, where they are chemisorbed or physisorbed; (3) adsorbed species on the surface react to form the solid crystal; (4) reaction products diffuse from the surface; and (5) reaction products are carried away in the flowing gas stream. The slowest of these five steps will determine the growth rate. Typically, the rate limiter is either step 2 or step 3, and these two situations are called mass transfer limited (or diffusion limited) and reaction rate limited, respectively.

Consider the transport of a single reactant to the growing surface. (Usually, one reactant is provided in excess for the growth of a binary compound, so that this assumption remains useful.) The flux of this species to the surface at a particular point is given by Henry's law:

$$j = h\left(N_g - N_0\right), \tag{3.1}$$

where:

N_g is the concentration of the reactant in the gas phase
N_0 is the concentration of the reactant at the surface
h is the gas phase mass transfer coefficient.

Suppose the reaction rate is linear; then

$$j = kN_0, \tag{3.2}$$

where k is the surface reaction rate constant. Usually, this rate is thermally activated so that

$$k = k_0 \exp(-E_a / kT), \tag{3.3}$$

where E_a is the activation energy for the process. Typical activation energies are in the range of 25–100 kcal mole^{-1} (1.1–4.3 eV molecule^{-1}).

Under steady-state conditions, the two fluxes above may be equated. Combining these equations, we can determine the growth rate as

$$g = \frac{j}{n} = \frac{N_g}{n}\left(\frac{hk}{h+k}\right), \tag{3.4}$$

where n is the number of atoms (or molecules) per unit volume in the growing crystal. For Si, $n = 5 \times 10^{22}$ cm^{-3} (Si), and for GaAs, $n = 2.2 \times 10^{22}$ cm^{-3} (Ga).

At low temperatures, $k \ll h$, so that

$$g \approx \frac{kN_g}{n}. \tag{3.5}$$

This is referred to as reaction-rate-limited growth. Under these conditions, the growth rate is a strong function of temperature. Here, the reaction rate is sensitive to the surface conditions and the growth rate depends on the orientation of the crystal substrate. This can result in faceted growth, which is usually undesirable, but can be advantageous in some situations, such as the implementation of epitaxial lateral overgrowth (ELO).

At high temperatures, $h \ll k$, so that

$$g \approx \frac{hN_g}{n}. \tag{3.6}$$

This situation is known as mass-transfer-limited growth (or diffusion-limited growth). Under mass-transfer-limited conditions, the growth rate is independent of the crystal orientation and nearly independent of temperature. There is a slight temperature variation (with an activation energy of 3–8 kcal mole^{-1}) due to the temperature dependence of the diffusivity. Usually, VPE reactors are designed to operate in the mass-transfer-limited regime. However, this is not always possible due to constraints imposed by the substrate or source chemicals.

Thermodynamic considerations may also be important in determining the growth rates in the mass-transfer-limited regime. This is illustrated in Figure 3.1, which shows the general behavior for the cases of the (a) endothermic and (b) exothermic processes. For the endothermic process, which involves a positive heat of reaction, the growth rate increases monotonically with increasing temperature. In the reaction-rate-limited regime, the activation energy is large, typically 25–100 kcal mole^{-1}, whereas in the mass-transfer-limited region, there is only a slight variation of the growth rate with temperature (3–8 kcal mole^{-1}). In contrast, for the exothermic process, which is favored at lower temperatures, the growth rate decreases rapidly with increasing temperature under conditions of mass-transfer-limited growth. Many processes of interest for heteroepitaxy are either endothermic (chloride VPE of Si) or pyrolytic (hydride VPE of Si or SiC, MOVPE) in nature, and so display the general behavior shown in Figure 3.1a.

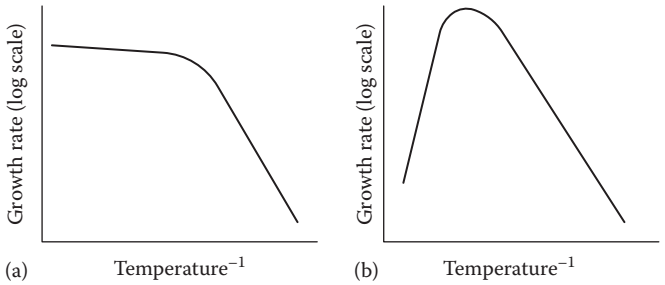


FIGURE 3.1
Growth rate (log scale) vs. reciprocal of temperature for (a) an endothermic process and (b) an exothermic process.

3.2.2 Hydrodynamic Considerations

The simplified model of Section 3.2.1 fails to reveal many of the details associated with the fluid dynamics of vapor phase epitaxy. Nor does it give guidance in the determination of the mass transfer coefficient. However, in the typical case of laminar flow, simple analytical solutions exist for the determination of the growth rate under mass-transfer-limited conditions.

The nature of the gas flow in an epitaxial reactor can be understood based on a study of gas flow in simple pipes.[1] This behavior may be characterized by the unitless Reynold's number, given by

$$N_R = \frac{dv\rho}{\mu}, \tag{3.7}$$

where:

- d is the pipe diameter
- v is the gas velocity in the pipe
- μ is the absolute viscosity
- ρ is the gas density

The viscosity of hydrogen, the most commonly used carrier gas, varies from about 200×10^{-6} dyn cm^{-1} s^{-1} (200 μP) at 700°C to about 250×10^{-6} dyn cm^{-1} s^{-1} (250 μP) at 1200°C. The density of hydrogen at atmospheric pressure varies from about 2.5×10^{-5} g cm^{-3} at 700°C to 1.65×10^{-5} g cm^{-3} at 1200°C. Empirically, it is found that the transition from laminar flow occurs with the Reynold's number in the range of $2000 < N_R < 3000$, and typical epitaxial reactors operate under laminar flow conditions, with $N_R \approx 30$.

Consider the idealized reactor shown in Figure 3.2, with a recessed susceptor and a constant cross-sectional area. The reactor height is h. Suppose that a single reactant contributes to the growth, and the concentration of this reactant is N_g at the entrance to the reactor. The growth is assumed to take place under mass-transfer-limited conditions so that $N_0 \approx 0$.

The growth rate will be proportional to the flux of reactant species arriving at the substrate surface. If this flux is controlled by diffusion, then

$$j = -D\frac{\partial N}{\partial y}, \tag{3.8}$$

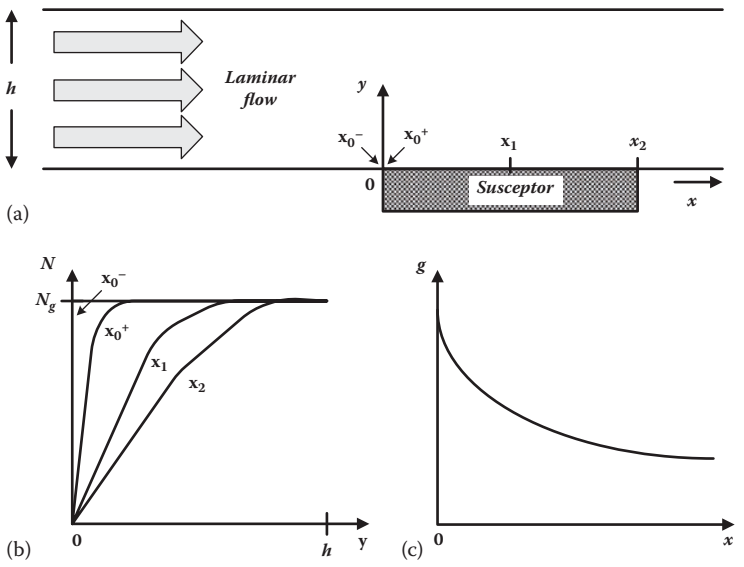


FIGURE 3.2
(a–c) Concentration of reactants in a horizontal reactor, and the growth rate. (Adapted from S. K. Ghandhi, *VLSI Fabrication Principles*, 2nd ed. [New York: John Wiley & Sons, 1994]. Used with permission.)

where:

- N is the actual concentration of the reactant in the gas phase at a point above the susceptor
- D is the diffusivity of the reactant species in the carrier gas
- y is the distance from the substrate interface

If the gas velocity is assumed to be constant above the susceptor, then under steady-state conditions (with all time derivatives equal to zero) the two-dimensional (2-D) continuity equation for the reactant species is

$$\frac{\partial N}{\partial t} = D\frac{\partial^2 N}{\partial x^2} + D\frac{\partial^2 N}{\partial y^2} - v\frac{\partial N}{\partial x} = 0, \tag{3.9}$$

where v is the gas velocity and x is the horizontal distance from the edge of the wafer. If diffusion is neglected in the flow direction, then

$$0 = D\frac{\partial^2 N}{\partial y^2} - v\frac{\partial N}{\partial x}. \tag{3.10}$$

The boundary conditions are

$$\begin{aligned} N &= N_g; && x = 0;\ 0 < y < h; \\ N &= 0; && 0 < x,\ y = 0; \\ \frac{\partial N}{\partial y} &= 0; && 0 < x,\ y = h. \end{aligned} \tag{3.11}$$

Solving, the flux of reactant is found to be

$$j = -\frac{2DN_g}{h}\sum_{r=0}^{\infty}\exp\left(\frac{-\pi^2 Dx(2r+1)^2}{4vh^2}\right). \tag{3.12}$$

If the flux of reactant species is envisioned to flow through a "diffusion boundary layer" in which its concentration varies linearly from N_g to 0, then this boundary layer thickness is given by

$$\delta_D(x) = \frac{h}{2}\left[\sum_{r=0}^{\infty}\exp\left(\frac{-\pi^2 Dx(2r+1)^2}{4vh^2}\right)\right]^{-1}. \tag{3.13}$$

Under the simplifying assumption that $(h^2v/\pi D)>x$, we obtain

$$j \approx -N_g\sqrt{\frac{Dv}{\pi x}}, \tag{3.14}$$

so that the diffusion boundary layer thickness is given by

$$\delta_D(x) \approx \sqrt{\frac{\pi Dx}{v}}, \text{ for } \delta_D(x) \le h, \tag{3.15}$$

and the growth rate may be estimated from

$$g \approx \frac{N_g D}{n\delta_D}. \tag{3.16}$$

It is apparent from this analysis that the achievement of uniform growth over a large wafer requires a uniform diffusion boundary layer thickness. This can be achieved by tilting the susceptor, so that the cross section of the reactor decreases with x. Sometimes, this is done in conjunction with adjustments to the total flow and reactor pressure in order to achieve the desired result.

The increased performance of computers has enabled detailed numerical computations of the mass and heat flows in epitaxial reactors. These calculations take into account continuity, conservation of momentum (Navier–Stokes equations), conservation of energy, and conservation of mass of diffusing species. Thus, the growth rate and uniformity (in thickness and composition) can be predicted for a reactor design before it is built and tested. Nonetheless, the simple boundary layer picture allows one to construct a starting point for the reactor design, which can then be fine-tuned by the use of computation.

3.2.3 Vapor Phase Epitaxial Reactors

A VPE reactor comprises a gas delivery system, a reaction chamber, and an effluent handling system; a basic setup is shown schematically in Figure 3.3. Welded stainless steel construction with metal gasket fittings is used to achieve the necessary leak integrity.

Typically, ultra-high-purity (UHP) H_2 with 7N purity (seven nines purity, or 99.99999% pure) is used as the carrier gas in a VPE reactor. This gas may be purchased in UHP

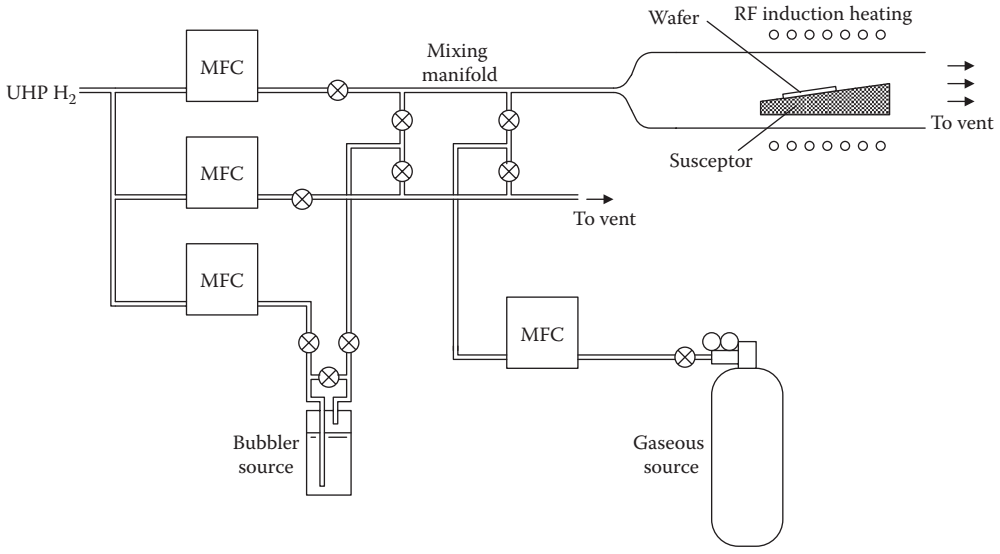


FIGURE 3.3
VPE reactor.

form, or commercial-grade hydrogen (3N5, or 99.95% pure) may be purified by diffusion through a palladium-silver membrane or by the use of a purifying resin. Other carrier gases may also be used, such as UHP N_2, He, or Ar. Although these gases preclude the use of a palladium cell, resin purifiers are available for use with them.

The source chemicals may be gaseous, liquid, or solid. Gas sources (such as SiH_4 and AsH_3) may be obtained in pure form, or diluted in hydrogen, in high-pressure cylinders. Liquid sources (such as SiH_2Cl_2 or TMGa) are typically obtained in UHP form, in stainless steel bubblers. Solid sources (such as TEIn) may also be obtained in bubbler vessels. However, it is difficult to obtain good run-to-run repeatability with these. Occasionally, vapor phase sources are created *in situ*, as in the case of GaCl, which was used as the Ga source in the hydride and halide epitaxial processes for GaAs.

For a gaseous source, the flow is metered precisely by an electronic mass flow controller (MFC). The MFC has a built-in heated capillary. The measurement of the temperature difference across the capillary allows the determination of the mass flow with a precision of ±0.5%. MFCs have built-in closed-loop control systems and metering valves, so they can maintain the flow at a desired set point. Typically, MFCs are calibrated for use with pure H_2 or N_2. Hydrogen calibration is entirely adequate for a dopant gas that is diluted to a few 100 ppm in H_2. In the case of pure gaseous sources such as AsH_3, the MFC must be calibrated specifically for each source.

Liquid sources are transported to the reactor by a carrier gas, usually H_2, which is passed through a stainless steel bubbler arrangement like the one shown in Figure 3.4. The carrier gas is metered precisely by an MFC placed upstream of the bubbler. If the mass flow of the carrier gas is F_{H_2}, the vapor pressure of the liquid source is P_S, and the total pressure in the bubbler is P_{TOT}, then the mass flow of the source is given by

$$F_S = F_{H_2}\left(\frac{P_S}{P_{TOT} - P_S}\right). \tag{3.17}$$

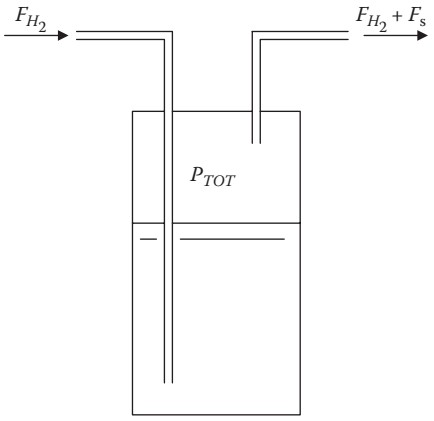


FIGURE 3.4
Liquid source bubbler.

This expression assumes that the carrier gas bubbles have sufficient residence time in the liquid to become saturated with its vapor, and closely approximates a real bubbler application. Typically, a three-valve arrangement is used around the bubbler so that the carrier gas can be made to bypass the bubbler when the source is not in use.

The vapor pressures of liquid sources are usually fit by the expression

$$\log_{10} P_S = A - \frac{B}{T}, \tag{3.18}$$

where:

P_S is the vapor pressure over the liquid (in torr)
T is the absolute temperature
A, B are empirical constants

For convenience, the temperature of a bubbler source is usually set to yield a source vapor pressure of 5–50 torr by means of a temperature-controlled bath. In some cases, the bubbler temperature must be kept above room temperature, which necessitates heating of the downstream lines to prevent condensation of the source.

The carrier gas and vapor phase sources are brought to a mixing manifold prior to their injection into the reactor chamber. In the mixing manifold, gas flows are switched between pressure-balanced vent and reactor lines. Thus, each flow may be stabilized to the vent line before being switched to the reactor. Precise pressure balancing, using a differential pressure transducer and a control valve, avoids unwanted flow transients.

VPE reactors are of the horizontal, vertical, or barrel types. The horizontal configuration is the simplest, and often used in research. The reaction chamber is a quartz tube, flanged at one end to facilitate loading and unloading of wafers. The substrate wafers are held by recesses in a graphite susceptor, which is usually tilted at a slope of 7°–10° to improve the thickness uniformity. The configuration is so named because of the horizontal gas flow in the tube.

The vertical reactor utilizes a vertical flow of gases, perpendicular to the surface of the wafers. Inherent in this geometry is a stagnation point at the center of the susceptor, where the gas velocity is zero. This tends to promote recirculation unless the susceptor is rotated at a high speed (>1000 rpm). Rotation also serves to improve the radial uniformity

of growth. In addition, the flow, pressure, and rotation rate must be optimized for radial uniformity. Sometimes, complex planetary rotation systems are employed as well.

The barrel reactor can handle many wafers in a single run and achieves high throughput. Wafers are held in shallow depressions within the steeply sloped susceptors, and the gas flow is nearly parallel to their surfaces. Thus, the barrel reactor geometry is similar to that of the horizontal reactor, but rotated 90°.

Heating of an epitaxial reactor may be accomplished by radio frequency (RF) induction, infrared lamps, or resistive heaters. In cold-wall reactors used for endothermic and pyrolytic reactions, RF induction or infrared lamps are typically used. Internal resistive heaters are occasionally used, but the materials must be chosen carefully to avoid metallic contamination. External resistive heaters avoid this problem but are only applicable to hot-wall reactors used for exothermic processes. Temperature control with ±2°C precision is normally adequate if the growth is mass transfer limited. Measurement of the temperature can be achieved using optical pyrometry, or for low-temperature reactors (<900°C), thermocouples may be embedded in the susceptor.

Susceptors are usually made from machined graphite. At temperatures above 1300°C, however, the hydrogen carrier gas will react with graphite and etch its surface. For this reason, SiC-coated susceptors are often employed in reactors intended for high-temperature operation. Coatings of SiC or pyrolytic BN are sometimes used in lower-temperature reactors as well, to avoid the outgassing affects associated with the porosity of uncoated graphite.

Epitaxial reactors may operate at atmospheric or reduced (~0.1 atm) pressure. Low-pressure operation reduces the surface coverage of adsorbed species, increasing their mobility and allowing high-quality growth at reduced temperatures (50°C–100°C lower than for atmospheric growth). Reduced pressure can also decrease the tendency for recirculation under otherwise similar conditions. In low-pressure reactors, pressure control is achieved using a mechanical vacuum pump and a butterfly valve. Pressure measurement may be by capacitance or optical manometers.

The effluent is typically treated by activated charcoal absorption units or liquid scrubbers before being vented to the outside. These systems require frequent service as well as expensive waste disposal.

3.2.4 Metalorganic Vapor Phase Epitaxy

The MOVPE process was developed in the late 1960s by Manesevit,[2–4] who first demonstrated its use for the epitaxy of Ga-V compounds. Subsequently, the process has been adapted to nearly all III-V and II-VI semiconductors, including the antimonides, arsenides, phosphides, nitrides, sulfides, selenides, and tellurides, and also ternary and quaternary alloys. MOVPE-grown material is of extremely high purity: this epitaxial method has produced the highest-purity InP produced by any method and GaAs, which is as pure as that grown by any technique. Specially designed reactors have also made possible the growth of very abrupt interfaces and multilayered structures of the type necessary for quantum layer devices such as laser diodes and high-speed transistors. These developments and the ease of scaling the MOVPE process to high throughput have made it important for commercial production, as well as laboratory research.

MOVPE goes by a number of names, including organometallic vapor phase epitaxy (OMVPE), metalorganic chemical vapor deposition (MOCVD), organometallic chemical vapor deposition (OMCVD), and occasionally organometallic epitaxy (OME). *Chemical vapor deposition* (CVD) is a more general term that applies to noncrystalline films; as such, the more specific term *vapor phase epitaxy* should be used to refer to epitaxy. OMVPE is

preferred by many researchers because it is consistent with the normal chemical nomenclature. MOVPE is the name chosen by the international conference and is used throughout this book. It is important to realize that these terms are used interchangeably in the literature, and are not meant to refer to different processes.

MOVPE is carried out in a reactor of the type shown schematically in Figure 3.3. Source chemicals are transported to the reactor by a carrier gas, where they react heterogeneously at the surface of a heated single-crystal substrate. In the growth of a binary semiconductor, one or both source chemicals may be metalorganic compounds. These are typically liquids at room temperature and are transported to the reactor by flow of carrier gas through a stainless steel bubbler. Gaseous sources may also be used; these are contained in high-pressure cylinders in either pure or diluted form. Ternary or quaternary alloys may be grown by introducing additional source chemicals. Changes in composition can be realized by ramping or switching the source flows. Layers may also be doped by the introduction of small concentrations of the appropriate sources.

MOVPE processes are pyrolytic in nature. As a consequence, cold-wall reactors are used almost exclusively. Also, the irreversible nature of MOVPE allows the growth of extremely abrupt interfaces.

The variety of metalorganic source chemicals has increased greatly over the years due to the success of the method and resulting market demand. In general, the sources are molecules of the type MR_n, where M represents a metal atom and R represents an organic radical. It is common practice to refer to the organic groups using M, E, NP, IP, NB, IB, TB, A, and Cp for methyl, ethyl, *n*-propyl, *i*-propyl, *n*-butyl, *i*-butyl, *t*-butyl, allyl, and cyclopentadienal, respectively. M, D, and T are used to denote mono-, di-, and tri-, respectively. Thus, TMGa represents trimethylgallium and DETe refers to diethyltelluride.

The metalorganic precursors are generally liquids at room temperature, contained in stainless steel bubbler vessels. A few are solid at room temperature, but these can be used with a bubbler arrangement as well. The melting points, boiling points, and vapor pressure parameters A and B are given in Tables 3.1 through 3.4 for sources of elements from columns II, III, V, and VI, respectively. As a general rule, the vapor pressures are highest for the lightest molecules.

The decomposition characteristics of the alkyl source molecules are determined in part by the strength of the metal–carbon bond. This bond energy determines the stability of the molecule with respect to decomposition by the removal of organic radicals (free radical homolysis). Therefore, it often determines the activation energy for reaction-rate-limited growth. In general, the metal–carbon bond strength decreases with the number of carbons bonded to the central carbon of the molecule (methyl > ethyl > isopropyl > tertiary butyl > allyl). In some situations, this means that a lower growth temperature may be used with an isopropyl source than with a methyl source. Bond strengths for some of the common alkyl precursors are provided in Table 3.5.

TABLE 3.1

Melting Points, Boiling Points, and Vapor Pressure Data for Metalorganic Sources of Column II Elements ($\log_{10}P(\text{torr}) = A - B/T$)

			Vapor Pressure		
Precursor	**Melting Point (°C)**	**Boiling Point (°C)**	***A***	***B* (K)**	***P* (torr) @ *T* (°C)**
DMZn	–42	46	7.802	1560	124 @ 0°C
DEZn	–28	118	8.280	2109	3.6 @ 0°C
DMCd	–2	106	7.764	1850	9.7 @ 0°C

TABLE 3.2

Melting Points, Boiling Points, and Vapor Pressure Data for Metalorganic Sources of Column III Elements ($\log_{10}P(\text{torr}) = A - B/T$))

			Vapor Pressure		
Precursor	**Melting Point (°C)**	**Boiling Point (°C)**	***A***	***B* (K)**	***P* (torr) @ *T* (°C)**
TMAl	15	126	8.224	2134.83	2.2 @ 0°C
TEAl	–52.5	186	10.784	3625	0.5 @ 55°C
TMGa	–15.8	55.8	8.501	1824	66 @ 0°C
TEGa	–82.5	143	9.172	2532	3.4 @ 20°C
DEGaCl	–7	—	8.78	2815	0.5 @ 60°C
TMIn	88	135.8	10.520	3014	0.3 @ 0°C
TEIn	–32	184			1.2 @ 40°C

TABLE 3.3

Melting Points, Boiling Points, and Vapor Pressure Data for Metalorganic Sources of Column V Elements ($\log_{10}P(\text{torr}) = A - B/T$)

			Vapor Pressure		
Precursor	**Melting Point (°C)**	**Boiling Point (°C)**	***A***	***B* (K)**	***P* (torr) @ *T* (°C)**
TMP	–84	38	7.7627	1518	381 @ 20°C
TEP	–88	127	8.035	2065	46.5 @ 50°C
TBP	4	54	7.586	1539	141 @ 10°C
TMAs	–87.3	50	7.3936	1456	238 @ 20°C
TEAs		140			5 @ 20°C
DMAs		36.3	7.532	1443	176 @ 0°C
DEAs		102	7.339	1680	40 @ 20°C
TBAs	–1	65	7.243	1509	32 @ –10°C
TMSb	–87.6	80.6	7.73	1709	48.9 @ 10°C
TESb	–98	160	7.90	2183	4 @ 25°C
TMBi	–107.7	110	7.628	1816	27 @ 20°C

TABLE 3.4

Melting Points, Boiling Points, and Vapor Pressure Data for Metalorganic Sources of Column VI Elements ($\log_{10}P(\text{torr}) = A - B/T$)

			Vapor Pressure		
Precursor	**Melting Point (°C)**	**Boiling Point (°C)**	***A***	***B* (K)**	***P* (torr) @ *T* (°C)**
DES	–100	91 ± 1	8.184	1907	47 @ 20°C
DTBS	—	149 ± 2	—	—	—
DMSe		57			
DESe	—	108	7.905	1924	7.2 @ 0°C
DMTe	–10	92 (82)	7.97	1865	65 @ 30°C
DMDTe		220	6.94	2200	0.26 @ 23°C
DETe	—	137	7.99	2093	7.1 @ 20°C
DIPTe	—	—	8.29	2309	2.6 @ 20°C

TABLE 3.5

Bond Strengths for Common Alkyl Precursor Molecules

Precursor	D_1 (kcal mole^{-1})	D_2 (kcal mole^{-1})	D_{ave} (kcal mole^{-1})
DMZn	51 (54)	47	42
DMCd	53	46	33
TMAl	65		66, 61
TEAl			58
TMGa	59.5	35.4	59
TEGa			57
TMIn	47		
TMP			66, 63
TMAs	62.8		55
TMSb	57	57	52, 47
DES			65

Generally, the alkyls of column II and column III elements are Lewis acids (electron acceptors), whereas the alkyls of column V and column VI atoms are Lewis bases (electron donors). It is possible for a gas phase reaction to occur between alkyls with Lewis acid and Lewis base character, resulting in an adduct. If the adduct so produced is a low-vapor-pressure molecule, it may not contribute to epitaxial growth, and in fact, it may give rise to fouling of the reactor. Such a parasitic reaction is highly undesirable. On the other hand, some adducts can contribute to growth, and these (such as TMIn-TEP) are sometimes used intentionally as sources.

3.3 Molecular Beam Epitaxy

MBE is a UHV technique that involves the impingement of atomic or molecular beams onto a heated single-crystal substrate where the epitaxial layers grow.[5] The source beams originate from Knudsen evaporation cells or gas source crackers. These can be turned on and off very abruptly by shutters and valves, respectively, providing atomic layer abruptness.

Because MBE takes place in a UHV environment, it is possible to employ a number of *in situ* characterization tools based on electron or ion beams. These provide the crystal grower with immediate feedback, and improved control of the growth process.

MBE has been developed to the point where nearly every semiconductor of interest may be grown using the technique, including III-V and II-VI semiconductors; Si, Ge, and $Si_{1-x}Ge_x$ alloys; and SiC and $Si_{1-x-y}Ge_xC$ alloys. However, III-phosphides are difficult to grow by MBE, and alloys involving As and P are especially challenging. Other drawbacks of MBE are the initial high cost and maintenance requirements of the UHV system and low growth rates. These drawbacks are offset to a large extent by the precise control and *in situ* characterization so that MBE is used extensively for commercial device production at this time.

An MBE reactor involves a number of source cells arranged radially in front of a heated substrate holder, as shown in Figure 3.5. The source cells supply all atoms necessary for the growth and doping of the required semiconductor layers; six or more cells may be required. The simplest type of source cell is a thermal evaporator (Knudsen cell), but other more

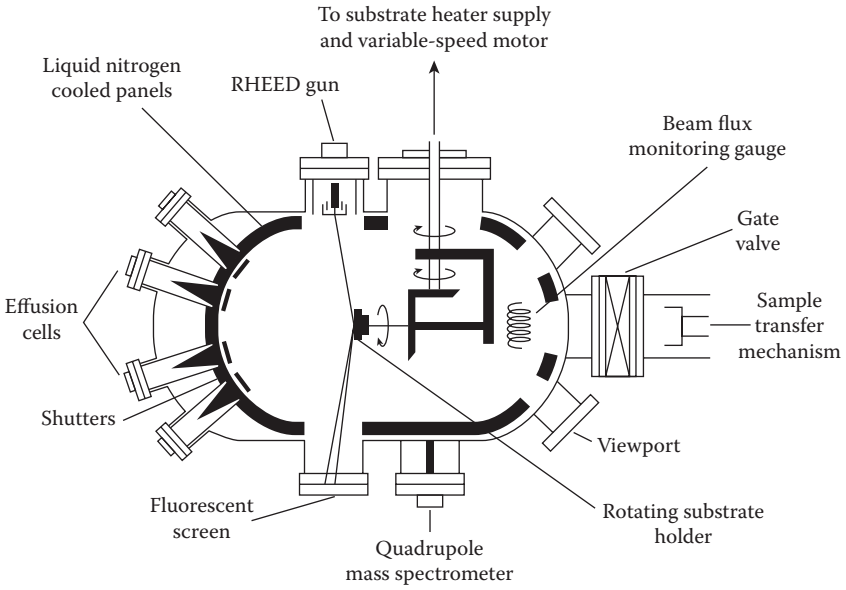


FIGURE 3.5
MBE reactor. (Reprinted from M. Henini, *Thin Solid Films*, 306, 331 [1997]. With permission. Copyright 1997, Elsevier.)

elaborate schemes have been developed for some atoms. A basic requirement for MBE growth is line-of-sight source impingement. This means that the evaporated source atoms must have mean free paths greater than the source-to-substrate distance, which is typically 5–30 cm. This requirement places an upper limit on the operating pressure for an MBE reactor.

The mean free path for an evaporated particle (atom or molecule) may be estimated if it is assumed that all other particles in the system are at rest. Suppose the evaporated particle is moving at a velocity c, and all particles have a round cross section with diameter σ. Two particles that pass at a distance of σ or less will collide. Therefore, each particle can be considered to have a collision cross section of $\pi\sigma^2$, and the collision volume swept out by a particle in time dt is $\pi\sigma^2 cdt$. If N is the volume concentration of particles, then the collision frequency will be

$$f = N\pi\sigma^2 c, \tag{3.19}$$

and the mean free path will be

$$\lambda = \frac{c}{f} = \left(N\pi\sigma^2\right)^{-1}. \tag{3.20}$$

A more accurate calculation of the mean free path may be made assuming that all of the particles are in motion. Based on this, the mean free path for an evaporated particle is

$$\lambda = \left(N\pi\sigma^2\sqrt{2}\right)^{-1} = \frac{kT}{\sqrt{2}\pi\sigma^2 P}, \tag{3.21}$$

where P is the pressure.

Typical values of the cross section diameter σ range from 2 to 5 Å, so that the mean free path is about 10^3 cm at a pressure of 10^{-5} torr. This pressure therefore represents an approximate upper limit for the system pressure during growth, if the beam nature of the sources is to be maintained.

The requirement on the base pressure is considerably more stringent, and is set by purity requirements. If the grown films are to have no more than 10^{-5} (10 ppm) contaminants, then the base pressure should be no more than 10^{-10} torr. Achievement of the necessary UHV requires the use of a stainless steel chamber with metal gaskets. The system must be load locked, so that it is opened to the atmosphere only for maintenance. Any exposure of the chamber to air must be followed by a long bakeout to remove adsorbed contaminants. During growth, the chamber walls must be cooled to cryogenic temperatures by means of a liquid nitrogen shroud, in order to further reduce evaporation from this large surface area.

Growth of pure layers by MBE also requires the use of oil-free pumping in the UHV system. Cryogenic sorption pumps, titanium sublimation ion pumps, and turbomolecular pumps are used for this reason.

The simplest source cells are thermal evaporators, called effusion cells or Knudsen cells. High-purity elemental sources are used, and one cell is needed for each element. Typically, the effusion cells are made of pyrolytic boron nitride with tantalum heat shields. The source temperatures are maintained precisely (±0.1°C) to control the flux of evaporating atoms. Due to the inability to rapidly ramp up or down the cell temperature, a shutter is used to turn each beam on and off.

The flux of atoms from such an effusion cell may be calculated using the kinetic theory of gases.[6] From this treatment, it can be shown that the evaporation rate from a surface area A_e is given by

$$\frac{dN_e}{dt} = \frac{A_e P}{\sqrt{2\pi kTm}}, \tag{3.22}$$

where:

P is the equilibrium vapor pressure of the source at the effusion cell temperature T
m is the mass of the evaporant

In terms of the molecular mass of the species, M, the effusion rate is

$$\frac{dN_e}{dt} = \frac{A_e P}{\sqrt{2\pi kTM / N_A}}, \tag{3.23}$$

where N_A is Avogadro's number (6.022×10^{23} mole^{-1}). Simplifying,

$$\frac{dN_e}{dt} = 3.51 \times 10^{22} \frac{A_e P}{\sqrt{MT}} \text{ molecules s}^{-1}, \tag{3.24}$$

where P is the pressure in torr. Because the equilibrium vapor pressure P varies exponentially with temperature, the effusion cell temperature must be controlled to within ±0.1°C in order to keep the effusion rate within a ±1% tolerance.

The flux of evaporant arriving at the substrate surface can be calculated from the evaporation rate at the effusion cell by

$$j = \frac{\cos\theta}{\pi l^2}\frac{dN_e}{dt}$$

$$= 1.117\times10^{22}\frac{A_e P\cos\theta}{l^2\sqrt{MT}}\ \text{molecules cm}^{-2}\ \text{s}^{-1}, \tag{3.25}$$

where:

l is the distance from the effusion cell to the substrate
θ is the angle between the beam axis and the normal to the substrate

The model outlined above assumes a full effusion cell so that evaporation occurs at its mouth. In practice, the cell depletes with time, and this causes a falloff of the impingement rate, and a change in the beam profile,[7] at the substrate. This effect can be mitigated to some extent by the use of tapered effusion cells.

Usually, the evaporation crucibles have a 1 cm^2 evaporation surface and are located 5–20 cm from the substrate. Typical source pressures are 10^{-3}–10^{-2} torr, resulting in the delivery of 10^{15}–10^{16} molecules cm^{-2} s^{-1}. This corresponds to a growth rate on the order of one monolayer per second, assuming a unity sticking coefficient for the impinging atoms.

Thermal effusion sources are switched on and off by means of pneumatically controlled shutters. A problem associated with this scheme is the change in thermal loading on the cell upon opening or closing the shutter. This causes unwanted temperature transients in the cell, which result in rather large variations (up to 50%) in the beam flux immediately after the shutter is opened.

Another disadvantage of thermal effusion sources is the inability to ramp the beam flux rapidly with time. Here, the limitation is due to the thermal mass of the effusion cell. This places an upper limit on the rate at which the composition may be ramped in a ternary or quaternary alloy. Whereas this restriction is important in the growth of multilayered or rapidly graded device structures, it is usually not a problem in graded buffer layers. In gas source MBE (GSMBE), the sources are controlled using MFCs that allow much more rapid ramping.

Electron beam evaporation sources have been used with some elements such as Si. Here, the elemental source is contained in a water-cooled crucible and evaporation occurs locally at the surface by the impingement of an electron beam. On and off control can be achieved by blanking of the electron beam. Scanning of the beam allows the realization of an extended-area source with characteristics similar to those of the thermal effusion cells.

3.4 Silicon, Germanium, and $Si_{1-x}Ge_x$ Alloys

Si, Ge, and their alloys may be grown by either VPE or MBE. Si(001) substrates are used almost exclusively for the epitaxy of these materials. Therefore, the *in situ* removal of the native oxide is a critical step prior to epitaxy. In the case of MBE, this can be achieved by flashing to a temperature up to 1200°C in the high vacuum. Prior to VPE growth, the oxide layer can be removed by a bakeout in hydrogen.

A number of sources can be used for Si VPE, including silicon tetrachloride ($SiCl_4$), trichlorosilane ($SiHCl_3$), dichlorosilane (SiH_2Cl_2), and silane (SiH_4); however, only dichlorosilane and silane are in common use at this time. This dichlorosilane process is heterogeneous (it requires two molecules of $SiCl_2$) and surface catalyzed (it occurs

only in the presence of the silicon surface). It is also reversible and is accompanied by etch-back and autodoping processes, whereby atoms from the grown crystal are etched and returned to the gas phase. These processes are undesirable in multilayered epitaxial device structures, because they compromise the abruptness of heterojunctions and also lead to nonideal doping profiles in p-n junctions. However, they can be suppressed by a reduction of the growth temperature.

The silane process is irreversible due to the absence of chlorine. Compared with the chlorosilanes, SiH_4 epitaxy can be carried out at a lower temperature but is extremely sensitive to oxidizing impurities. Silane epitaxy therefore mandates the use of load locks and careful bakeout procedures to avoid the formation of silica dust, which is detrimental to layer morphology. A unique aspect of the silane process is that homogeneous, gas phase nucleation is possible with this source.[8] The dusting that results from homogeneous nucleation can also deteriorate layer quality. However, this problem can be minimized by the use of low-pressure, high gas velocities, and reduced temperature.

The VPE growth of Ge has been achieved using a number of halogenic sources,[9,10,11] including germanium tetrabromide ($GeBr_4$), germanium tetrachloride ($GeCl_4$), and germanium diiodide (GeI_2), as well as the hydride, germane (GeH_4). The halide processes are reversible, and therefore accompanied by undesirable autodoping effects, which limit the abruptness of junctions. In addition, the iodide process is complex, requiring three separate temperature zones. For these reasons, the germane process is the preferred method of growth for germanium today and is used for the realization of Ge/Si and germanium-on-insulator (GOI).

$Si_{1-x}Ge_x$ epitaxy may be carried out using a mixture of silicon and germanium sources in the vapor phase. The gas phase mole fraction is used to control the resulting solid phase mole fraction x. Practical systems for $Si_{1-x}Ge_x$ VPE utilize $SiH_2Cl_2 + GeH_4$ or any combination of the sources SiH_4, Si_2H_6, GeH_4, and Ge_2H_6. In the case of $SiH_2Cl_2 + GeH_4$, the solid composition x depends on the ratio of the gas source flows by[12]

$$\frac{x^2}{1-x} = 2.66 \frac{X_{GeH_4}}{X_{SiH_2Cl_2}}. \tag{3.26}$$

Selective growth of $Si_{1-x}Ge_x$ may be achieved by the use of $SiH_2Cl_2 + GeH_4 + HCl$; growth proceeds on bare silicon surfaces but not on dielectric films such as SiO_2 or silicon nitride. This can be utilized in patterned or nanoheteroepitaxial growth schemes. Commercial $Si_{1-x}Ge_x$ VPE reactors provide for the use of either combination of sources, to allow either nonselective (blanket) or selective growth. However, growth over a dielectric film is polycrystalline and should be properly referred to as chemical vapor deposition, not vapor phase epitaxy.

In addition to $Si_{1-x}Ge_x$, the carbon-containing alloys $Si_{1-y}C_y$ and $Si_{1-x-y}C_yGe_x$ are of interest for bandgap engineering of heteroepitaxial devices on Si wafers. These materials may be grown by the addition of a carbon precursor to the growth chemistry, and practical VPE systems employ monomethylsilane ($SiCH_6$) for this purpose. Due to the extremely low solubility of C in Si ($<10^{-6}$), all practical carbon-containing alloys are necessarily metastable[13] and must be grown at low temperatures.

UHV vapor phase epitaxy[14] has also been used to grow $Si_{1-x}Ge_x$ alloys, with a growth pressure of ~10^{-3} atm. Under UHV conditions, good-quality layers may be grown with a cold-wall reactor and the homogeneous nucleation is suppressed. Any combination of the sources SiH_4, Si_2H_6, GeH_4, and Ge_2H_6 may be used, but the use of Si_2H_6 has been reported to provide better surface morphology.

$Si_{1-x}Ge_x$ alloys across the entire compositional range may be grown by MBE using electron beam sources. Typical temperatures range from 500°C to 900°C. Usually, films with higher Ge content are grown at lower temperatures, keeping the growth temperature at approximately 60%–70% of the melting temperature. Temperature ramping may be employed during the growth of a graded layer. A unique aspect of MBE growth is the ability to grow $Si_{1-x}Ge_x$ films at very low temperatures[15] (300°C–400°C); the altered kinetics of lattice relaxation appears to enable the growth of layers with reduced threading dislocation densities. Kinetic models for lattice relaxation are described in Chapter 5.

In the fabrication of multilayered device structures, there is a tendency for three-dimensional (3-D) growth of $Si_{1-x}Ge_x$ on Si, except at low values of x. However, the islanding can be suppressed by reduction of the growth temperature (<600°C for MBE), thus allowing the growth of strained-layer superlattices.[16] Equilibrium modes of growth for SiGe/Si (111) and SiGe/Si (100) are discussed in Chapter 4.

3.5 Silicon Carbide

Epitaxial SiC may be grown using various combinations of precursors. The most commonly used silicon source is SiH_4,[17] but Si_2H_6[18] and $SiCl_4$[19] have also been used. The most popular carbon source is C_3H_8, but the sources C_2H_2,[20] CH_3Cl,[21] CH_4,[22] CCl_4, C_7H_8, and C_6H_{14} have been used as well. Some researchers have even demonstrated the growth of SiC from a single precursor. Sources of this type include CH_3SiCl_3[23] and $(CH_3)_2SiCl_2$.[22]

Usually SiC epitaxy is carried out in the system $SiH_4 + C_3H_8 + H_2$ in the temperature range of 1200°C–1800°C, with growth rates of 1–5 μm h^{-1}.[24] The quality of the epitaxial SiC is strongly dependent on the C/Si ratio in the gas phase. Typically, this ratio is 3:1, corresponding to a C_3H_8/SiH_4 ratio of 1:1, although this depends on the reactor.

SiC may also be grown by GSMBE using the sources $SiH_4 + C_2H_4$[25] or $Si_2H_6 + C_2H_4$.[26,27] Using the sources $SiH_4 + C_2H_4$, and 0.75 sccm flow of each, very low growth rates are obtained: 3 nm h^{-1} at 1000°C to ~50 nm h^{-1} at 1500°C. The addition of H_2 increases the growth rates dramatically (0.2 μm h^{-1} at 1500°C). The growth rate depends on both source flows because neither is in strong excess.

The most common substrate for heteroepitaxial growth of SiC is Si. Usually, Si (111) is used for the heteroepitaxy of 4H-SiC or 6H-SiC, due to the threefold symmetry of its surface. On the other hand, the cubic polytype 3C-SiC may be grown heteroepitaxially on Si (001).

Misoriented substrates are often employed for SiC epitaxy. For heteroepitaxy of 3C-SiC on Si (001), it is necessary to use misoriented substrates in order to eliminate inversion domain boundaries. For homoepitaxy of 6H-SiC (0001), the growth temperature may be lowered significantly (e.g., from 1800°C to 1500°C) by the use of substrates that are misoriented by a few degrees from the (0001) plane toward a $\langle 11\bar{2}0 \rangle$ direction. Growth on exact (0001) substrates at low temperatures is characterized by mixed 3C and 6H phases, but misorientation of the substrate by 1° or more toward the $\langle 11\bar{2}0 \rangle$ direction eliminates this problem and allows growth of single-phase 6H-SiC at 1500°C. This technique is called step-controlled epitaxy[28,29] and is now commonly employed for the fabrication of SiC devices.

3.6 III-Arsenides, III-Phosphides, and III-Antimonides

GaAs and the ternary $Al_xGa_{1-x}As$ may be grown by MOVPE or MBE, and commercial production of $Al_xGa_{1-x}As$ lasers is split between these two techniques. These materials are most often grown on GaAs substrates. However, heteroepitaxy on Si and InP substrates has been investigated extensively with the goal of integrating $Al_xGa_{1-x}As$ devices with those from these other material systems.

MOVPE growth is carried out with the sources TMGa + TMAl + AsH_3. Typically, a growth rate of ~10 μm h^{-1} is achieved using a mole fraction $X_{TMGa} = 10^{-4}$ and a growth temperature of 650°C. The growth rates for GaAs and AlAs are proportional to the respective organometallic source mole fractions. The V/III ratio is 5–30 for atmospheric pressure, but higher values may be used for reduced pressure growth.

Truly selective area growth of GaAs is possible using the source combination DEGaCl + AsH_3 and a SiO_2 mask. The GaAs grows where windows have been opened in the oxide, but there is no deposition on the oxide itself. Moreover, this approach can be extended to $Al_xGa_{1-x}As$ by using DEGaCl + DEAlCl + AsH_3.

In the case of MBE, elemental sources (7N Ga, 6N As, and 6N Al) are used in conventional Knudsen cells. A temperature of 550°C–600°C is used, with growth rates of 0.1–1 μm h^{-1}.

GaAs heteroepitaxy on Si (001) substrates raises a number of challenging problems. The growth mode is 3-D (Volmer–Weber), so a low-temperature nucleation layer must be used in order to obtain a smooth device layer. Inversion domain boundaries (also known as antiphase domain boundaries) are produced if on-axis substrates are used, but this problem can be eliminated by the use of Si substrates that are misoriented by 2°–4°. Inversion domain boundaries are considered in Chapter 4. The large lattice mismatch (~−4%) results in large threading dislocation densities (~10^8–10^9 cm^{-2}). Also, GaAs has about twice the thermal expansion coefficient of Si, so a large tensile strain is introduced in the GaAs during cooldown. This causes cracking in layers greater than about 4 μm thickness.

$In_xGa_{1-x}As$ is an important material for the channel regions of high-electron-mobility transistors (HEMTs) and also detectors for fiber optic communication systems operating in the range 1.3–1.55 μm. These materials can be grown by MOVPE using TEIn + TMGa + AsH_3 or TMIn + TMGa + AsH_3.[30,31] The ethyl source participates in a parasitic reaction with arsine unless the growth pressure is reduced to ~0.1 atm. This problem is eliminated with TMIn so that high-quality material is obtained with atmospheric growth. Usually, $In_xGa_{1-x}As$ is grown with the methyl sources at 650°C with a growth rate of ~3 μm h^{-1}. Other alloys involving Al can be grown by the addition of TMAl.

$In_xGa_{1-x}As$ is usually grown heteroepitaxially on InP or GaAs substrates. This material can be lattice matched to InP with $x = 0.53$. However, small processing variations in the composition result in the introduction of a large density of threading dislocations. $In_xGa_{1-x}As$ grown on GaAs is only lattice matched with $x = 0$. Linearly graded buffer layers (such as $In_xGa_{1-x}P$)[32] are often employed to transition from the lattice constant of GaAs to that of the $In_xGa_{1-x}As$ device layer. Here, the threading dislocation is found to be proportional to the grading coefficient, and in practical layers, dislocation densities as low as ~10^5 cm^{-2} may be obtained.

$Al_xIn_yGa_{1-x-y}P$ is an important material for high-brightness visible light-emitting diodes (LEDs) such as those used in street signs, traffic lights, and automotive applications, and for solar cells. This material can be lattice matched to GaAs and is usually grown heteroepitaxially on this substrate. This material is grown by MOVPE in the range 600°C–650°C.

With methyl sources in the combination TMAl + TMIn + TMGa + PH_3, growth pressures up to 1 atm can be used without parasitic reactions.

If the $Al_xIn_yGa_{1-x-y}P$ material is constrained to lattice match the GaAs substrate, then the indium content must be fixed at approximately $y = 0.5$. The material compositions that match the lattice constant of GaAs may therefore be written as $(Al_xGa_{1-x})_{0.5}In_{0.5}P$. The energy gap of this material lattice matched to GaAs is given by

$$E_g = 1.91 + 0.61x. \tag{3.27}$$

Even though the active layers of an $Al_xIn_yGa_{1-x-y}P$ LED may be lattice matched to the GaAs substrate, commercial high-brightness devices make use of highly mismatched heteroepitaxial GaP window layers, which spread the current of the top contact and greatly improve the device efficiency. A GaP substrate could serve to further reduce substrate absorption, but this approach is not used due to the high threading dislocation density it would produce in the active layers of the LED.

$In_{0.x}Ga_{1-x}P$ may also be used for visible LEDs in the orange and red portion of the spectrum. These devices are usually fabricated by heteroepitaxy on GaAs (001) substrates. However, $In_{0.5}Ga_{0.5}P$ LEDs have also been demonstrated on Si (001) substrates.[33] These devices were grown using GaAs buffer layers and exhibited stable output at 660 nm despite the very high threading dislocation density (~10^7 cm^{-2}).

The III-antimonides are of interest for applications as barrier layers in HEMTs,[34] focal-plane detector arrays in the 3–5 μm atmospheric window, and the fabrication of thermophotovoltaic devices.[35] These materials include InSb, AlSb, GaSb, and their alloys, and may be grown by MOVPE[36] or MBE.[37] InSb and GaSb substrates have relatively high threading dislocation densities, so GaAs[38,39] or InP substrates are usually used. Typically, a growth rate of ~2.5 μm h^{-1} is obtained at 600°C using the methyl sources TMGa, TMAl, TMIn, and TMSb.

For InSb grown on GaAs (001), the extremely large lattice mismatch strain ($|f| \sim 15\%$) gives rise to very high misfit and threading dislocation densities. This results in degradation of the carrier mobility near the heterointerfaces, which is associated with the misfit dislocations or a high-density tangle of threading dislocations, so that improved mobility is obtained away from the interface.

3.7 III-Nitrides

GaN, InN, AlN, and their alloys exist in the wurtzite structure and are grown almost exclusively on sapphire (0001) or hexagonal 6H-SiC (0001) substrates. However, growth on semipolar and nonpolar orientations of sapphire, as well as Si(111) substrates,[40,41] has been investigated as well.

The III-nitrides grow in a 3-D island mode on sapphire (0001) or 6H-SiC (0001) substrates. Therefore, low-temperature AlN nucleation layers[42,43,44] are commonly used to achieve smooth layers free from large columnar islands. Typically, the AlN nucleation layer is grown at a temperature of 450°C–550°C, whereas single-crystal AlN is grown by MOVPE at ~1000°C. The low-temperature AlN grows as an amorphous layer but crystallizes during a subsequent heat treatment. The success of the recrystallization process depends on a thin nucleation layer, so typically this thickness is 50 nm or less. Low-temperature GaN

nucleation layers,[45,46] have also been used, with similar improvements in the overgrown GaN material. A discussion of low-temperature nucleation layers is given in Chapter 4.

For the heteroepitaxy of GaN on sapphire, nitridation of the sapphire surface prior to growth is a critical step for the attainment of good crystal quality.[47] This step serves to replace O atoms by N to form a thin AlN layer. The change in the nucleation surface improves the final threading dislocation density in the overgrown material by a factor of 1/50.

The growth of III-nitrides on highly mismatched substrates gives rise to very high threading dislocation densities in the material. For GaN, the lattice mismatch is ~16% with sapphire (0001), which has the rhombohedral crystal structure with $a = 4.7592$ Å and $c = 12.9916$ Å.[48] Therefore epitaxial lateral growth (ELO) has been applied to obtain material with low threading dislocation densities for LEDs and laser diodes. ELO, and the related technique of pendeo-epitaxy, are described in detail in Chapter 8. Both of these approaches depend on the large lateral-to-vertical growth rate ratio obtained using MOVPE with the [0001] growth direction.

The III-nitrides must be grown at relatively high temperatures by either MBE or MOVPE (1000°C–1100°C for MOVPE of GaN), so considerable thermal strain is introduced during temperature changes. Sapphire has a larger coefficient of thermal expansion than GaN at room temperature. However, this situation reverses at higher temperatures, so that a tensile strain is introduced in the GaN during the cooldown process. Thermal strain and cracking are described in Chapter 5.

3.7.1 Vapor Phase Epitaxial Growth of III-Nitrides

In early work, Maruska and Tietjen[85] demonstrated the VPE of GaN in the Ga + HCl + NH_3 system. This approach has been replaced by MOVPE, using TMGa + NH_3, TMAl + NH_3, or TMIn + NH_3. Ternary or quaternary layers may be grown by using the appropriate combinations of the metalorganic sources. Typically, a high V/III ratio is used, so the alkyl flows determine the growth rate and composition of the epitaxial layer. A relatively high substrate temperature must be used for the MOVPE growth of any of these III-Nitrides, due to the thermal stability of ammonia. Two problems in the MOVPE growth of III-nitrides involve the prereactions between NH_3 and the metalorganic sources and the thermal convection of NH_3, which can disturb the laminar flow conditions and result in recirculation. Several novel reactor designs have been developed to cope with these problems. Akasaki et al.[49] demonstrated the growth of $Al_xGa_{1-x}N$ on sapphire (0001) using a "dual-flow" reactor design, shown in Figure 3.6. In this cold-wall, RF-heated reactor, the NH_3 and metalorganics are mixed a short distance before the reactor, and are directed at the slanted susceptor with high velocity (110 cm s^{-1}) to minimize parasitic reactions.

Nakamura et al.[50] designed a vertical, resistance-heated "two-flow" reactor for the growth of GaN, as shown in Figures 3.7 and 3.8. In this design, the main flow, comprising 5 slm NH_3, 54 μmol min^{-1} TMGa, and 1 slm H_2, was directed through a quartz nozzle in a direction parallel to the substrate surface. A subflow comprising 10 slm H_2 and 10 slm N_2 was directed perpendicular to the substrate surface, to bring the reactant gases in contact with the substrate. In their work, it was found that the perpendicular subflow was necessary in order to obtain a continuous film of GaN on a sapphire (0001) substrate at a growth temperature of 1000°C. Without it, only isolated islands of GaN were produced.

In the counterflow reactor of Lee et al.,[51] illustrated in Figure 3.9, the NH_3 is introduced through a counterflow while the organometallics are introduced through a front flow, thereby minimizing parasitic reactions and preheating the NH_3. Mirror-smooth layers of GaN were grown on sapphire (0001) at 1070°C using a total flow of 6 slm.

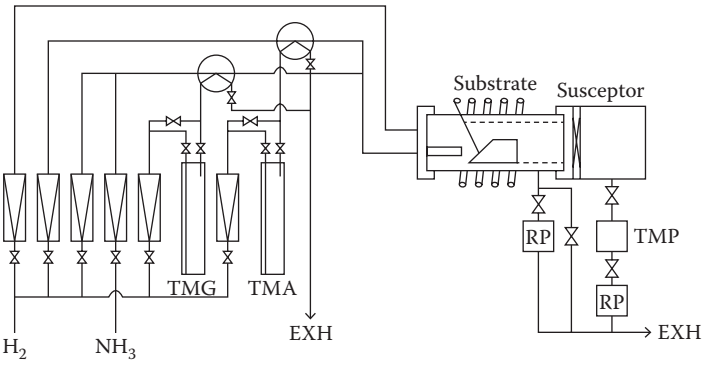


FIGURE 3.6
Dual-flow MOVPE reactor for the growth of $Al_xGa_{1-x}N$. (Reprinted from I. Akasaki et al., *J. Cryst. Growth*, 98, 209 [1989]. With permission. Copyright 1989, Elsevier.)

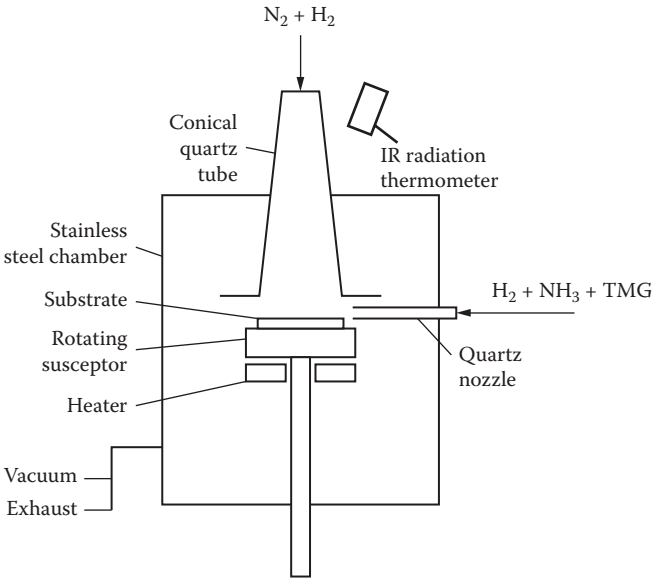


FIGURE 3.7
Two-flow MOVPE reactor for GaN growth developed by Nakamura et al.[50] (Reprinted from S. Nakamura et al., *Appl. Phys. Lett.*, 58, 2021 [1991]. With permission. Copyright 1991, American Institute of Physics.)

In many of the early studies of MOVPE growth of III-nitrides, high temperatures (>1000°C), high gas velocities (>1 m s^{-1}), and high ammonia flow rates (>5 slm) were required to obtain high-quality films. However, these conditions were deemed severe compared with those used in the growth of other III-V materials, leading to the development of other novel reactor designs using shrouding flows or NH_3 counterflow. In the two-flow cold-wall reactor of Nishida et al.[52] shown in Figure 3.10, an upper flow of H_2 was used to suppress convective effects associated with the NH_3. Mirror-smooth layers were obtained with reduced temperature (950°C) and reduced flows (1.5 slm of NH_3, TMGa, and H_2 as the carrier gas for the TMGa bubbler). An upper flow of 1 slm was used in their study.

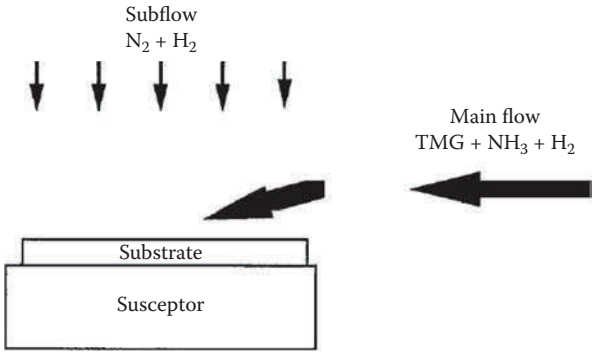


FIGURE 3.8
Details of the two-flow MOVPE reactor for GaN growth developed by Nakamura et al.[50] (Reprinted from S. Nakamura et al., *Appl. Phys. Lett.*, 58, 2021 [1991]. With permission. Copyright 1991, American Institute of Physics.)

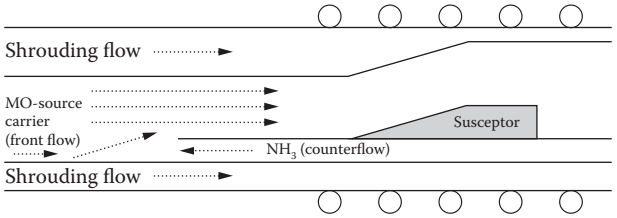


FIGURE 3.9
Counterflow MOVPE reactor for the growth of III-nitrides. (Reprinted from C.-R. Lee et al., *J. Cryst. Growth*, 182, 11 [1997]. With permission. Copyright 1997, Elsevier.)

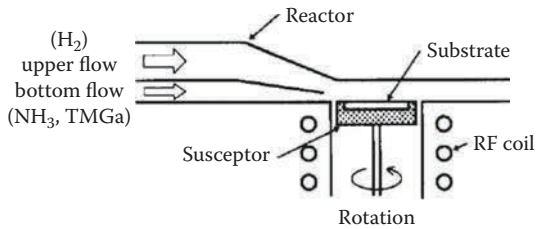


FIGURE 3.10
Two-flow GaN MOVPE reactor developed by Nishida et al.[52] (Reprinted from K. Nishida et al., *J. Cryst. Growth*, 170, 312 [1997]. With permission. Copyright 1997, Elsevier.)

3.7.2 Molecular Beam Epitaxy of III-Nitrides

MBE can also be used to grow the III-nitrides,[53] and has the advantage of allowing lower growth temperatures. Either RF plasma cells or compact electron cyclotron resonance (ECR) microwave plasma sources of nitrogen are employed.[54,55,56] The plasma source produces free radicals of nitrogen; without it, both NH_3 and N_2 are too stable to contribute to growth at typical temperatures employed for MBE growth. One such MBE system reported by Lin et al. is shown in Figure 3.11. In addition to the MBE chamber, this system has a UHV CVD system attached through a load lock, in which substrates may be cleaned in a hydrogen plasma prior to MBE growth. In such an MBE system the growth rate is limited

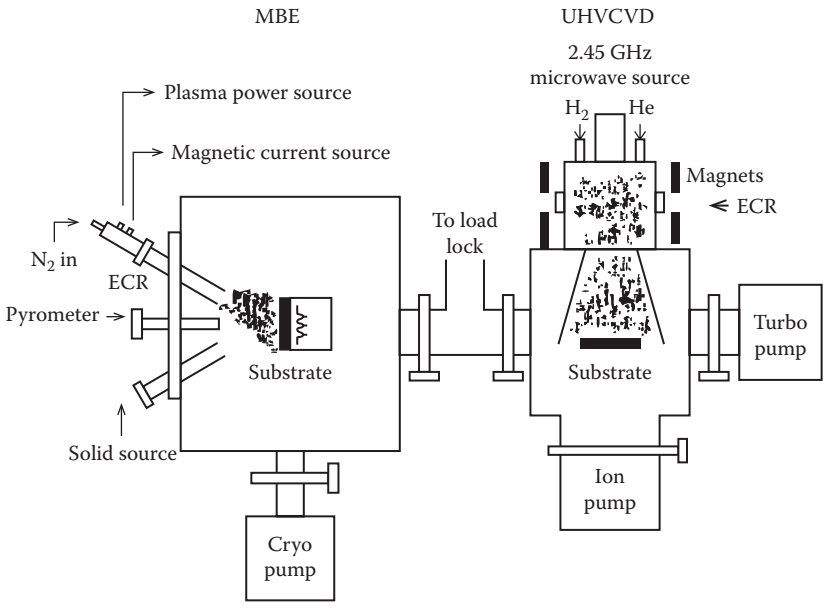


FIGURE 3.11
GaN MBE system utilizing an ECR microwave plasma source of nitrogen. (Reprinted from M. E. Lin et al., *J. Appl. Phys.*, 74, 5038 [1993]. With permission. Copyright 1993, American Institute of Physics.)

by the supply of active nitrogen from the plasma source, so that operation with a high Ga flux results in the formation of Ga droplets on the surface.

MBE-grown III-nitride materials may be doped n-type or p-type during growth through the use of appropriate dopants. Si, Ge, and Se have all been used as donor impurities, but Si is most commonly used, and electron concentrations as high as 10^{20} cm^{-3} have been achieved.[57] p-Type doping is invariably achieved using the impurity Mg. As-grown films are p-type, without the need for thermal annealing. Nitrogen-rich conditions are favorable for Mg incorporation, and concentrations exceeding 10^{20} cm^{-3} have been demonstrated.[58]

The dilute nitrides GaInNAs and GaInNAsSb have potential applications in optoelectronics for high-bit-rate communications systems, such as 1.2–1.6 μm lasers and optical amplifiers. These materials have been grown heteroepitaxially on GaAs substrates by MBE.[59] Due to the low solubility of N in these materials, they are susceptible to phase separation; therefore, metastable alloys must be grown at low temperatures (~425°C for MBE). The growth mode is SK (2-D growth of a wetting layer following by 3D island growth), but the 2-D to 3-D transition can be suppressed by the introduction of an Sb surfactant.

3.8 II-VI Semiconductors

The II-VI semiconductors include all combinations of Zn, Cd, Hg, and S, Se, and Te, and may be grown by MOVPE or MBE. These materials are usually grown heteroepitaxially on GaAs, InP, or Si. ZnSe substrates are available, but with relatively small area and high defect densities. However, device structures can be designed to be lattice matched to either GaAs or InP substrates. $Hg_{1-x}Cd_xTe$ is of great interest for infrared devices, and may also

be grown by MOVPE and MBE. Substrates such as CdTe, InSb, and even CdZnTe have been utilized, but GaAs or Si are used more commonly due to their larger area and better quality. ZnO is a direct bandgap semiconductor with $E_g = 3.37$eV and a large excition binding energy of 60 meV, so that room temperature excitonic UV emission should be possible.[60] This has led to great interest in the epitaxial growth of this material, which has been achieved by plasma-assisted MBE,[61] magnetron sputtering,[62] and pulsed laser deposition (PLD)[60] on a variety of substrates, including sapphire (0001), Si (001),[60] and CaF_2 (111).[61]

3.8.1 ZnSe and Its Alloys

ZnSe and its alloys must be grown at low temperatures in order to minimize the influence of native defects and their complexes. MOVPE growth at relatively low temperatures is possible using the hydride sources H_2Se, H_2S, and H_2Te. However, these hydrides give rise to gas phase prereactions with the metalorganics, degrading the layer quality and fouling the reactor. For this reason, the metalorganic sources such as DMSe, DES, and DETe are commonly used. These sources require an increase in the growth temperature unless photoirradiation from an ultraviolet lamp is employed. This photoassisted growth technique is complicated by the change of surface composition during heteroepitaxial growth. The mechanism of photoassisted growth appears to involve photogenerated carriers near the surface of the ZnSe, which promote the breaking of the alkyl source bonds. However, photoassisted carriers in a GaAs substrate do not participate in this process. Therefore, heteroepitaxial growth on GaAs substrates requires the growth of a high-temperature ZnSe buffer prior to the start of photoassisted epitaxy.

An important problem in the heteroepitaxy of wide-bandgap II-VI materials on GaAs substrates is the creation of stacking faults at the interface. It has been found that a single such defect can give rise to the rapid degradation and failure of an LED or laser diode. The nucleation of the stacking faults is related to the initial condition of the surface. In MBE growth, the formation of stacking faults can be suppressed by Zn stabilization (starting the Zn beam first).

3.8.2 HgCdTe

The ternary material $Hg_{1-x}Cd_xTe$, which is of great interest for infrared detectors in the 8–16 μm range, is usually grown on GaAs substrates by MOVPE[63,64] or MBE.[65,66] Occasionally, Si or sapphire[67] substrates have also been utilized. $Hg_{1-x}Cd_xTe$ exhibits a very large lattice mismatch (~14%) with GaAs substrates over the entire compositional range. As a consequence, it has been reported that the epitaxial relationship can be either CdTe[001]||GaAs[001] or CdTe[111]||GaAs[001].[68,69]

$Hg_{1-x}Cd_xTe$ is typically grown on GaAs substrates by MOVPE with the sources Hg + DMCd + DETe.[70] Typical growth temperatures range from 350°C to 420°C. It is found that these layers contain ~500 cm^{-2} hillocks, which may be associated with stacking faults at the interface. However, this problem can be prevented by the inclusion of a CdTe buffer layer. $Hg_{1-x}Cd_xTe$ can also be grown at a lower temperature (~175°C) using the source combination Hg + DMCd + DTBTe.[71]

In order to reduce the dislocation densities in $Hg_{1-x}Cd_xTe$ device layers grown on GaAs substrates, $Cd_{1-x}Zn_xTe$ buffer layers have been used, both the graded and constant composition types.[72] (Uniform and graded buffer layer approaches are discussed in Chapter 6.) Also, wide-bandgap barrier layers are used to reduce the interface recombination velocity.

3.8.3 ZnO

ZnO may be grown by plasma-enhanced MBE,[73–77] MOVPE,[78–81] magnetron sputtering,[62] or PLD[60,82,83] on a variety of substrates. Invariably, a low-temperature buffer layer is used to achieve smooth surface morphology and control the film orientation.[61,62,84]

MBE growth is conducted at a temperature of 550°C–600°C on a suitable buffer. Typically, the buffer layer is low-temperature ZnO deposited at a temperature of 300°C–350°C,[60,61] but occasionally other buffers have been used. Park et al.[74] demonstrated the epitaxy of smooth, high-quality ZnO on sapphire substrates using a low-temperature ZnO buffer on top of a Cr_2O_3 buffer created by the oxidation of an MBE-deposited chromium layer. MBE growth utilizes an RF plasma source[74] or an ECR microwave plasma source[76] for oxygen and a solid source for zinc.[84] Nominally, undoped films are invariably n-type. Hirano et al.[73] demonstrated that the use of flux-modulated MBE, in which the growth conditions are alternated between oxygen-rich and zinc-rich conditions, resulted in a lower residual electron concentration (9.7×10^{16} cm^{-2}) compared with standard oxygen-rich conditions (2.4×10^{17} cm^{-2}). p-Type ZnO has been reported using the dopants nitrogen,[77] phosphorus, and arsenic.

ZnO growth by MOVPE may be carried out in a broad temperature range, but typically 300°C–400°C is used. The commonly used sources are diethylzinc and tertiary-butanol (t-BuOH), with Ar as the carrier gas,[78] but these are not available in high purity. Dietylzinc and oxygen have also been used, but a parasitic gas phase reaction is problematic with this combination of sources. Ogata et al.[80,81] have demonstrated MOVPE growth of ZnO using diethylzinc and nitrous oxide (N_2O) at relatively high growth temperatures (600°C–700°C), but without gas phase reactions.

PLD may be conducted using a Nd:yttrium-aluminum-garnet (YAG) laser (355 nm, 5 Hz, 3.8 J cm^{-2}) to ablate a ZnO target, typically in a partial oxygen pressure of 1 torr,[60] in a system of the type shown in Figure 3.12. Under usual conditions it is necessary to use the oxygen ambient in order to obtain stoichiometric films. However, Yamaguchi et al.[82] have demonstrated the growth of stoichiometric films without an oxygen ambient by applying a bias voltage to a molybdenum grid placed between the target and substrate, as shown in Figure 3.12. The bias voltage is applied on the molybdenum grid with respect to the target. Low deposition temperatures (~200°C) may be used for PLD growth of ZnO, but it is often necessary to anneal the films at a higher temperature (~700°C) to obtain good crystallinity.

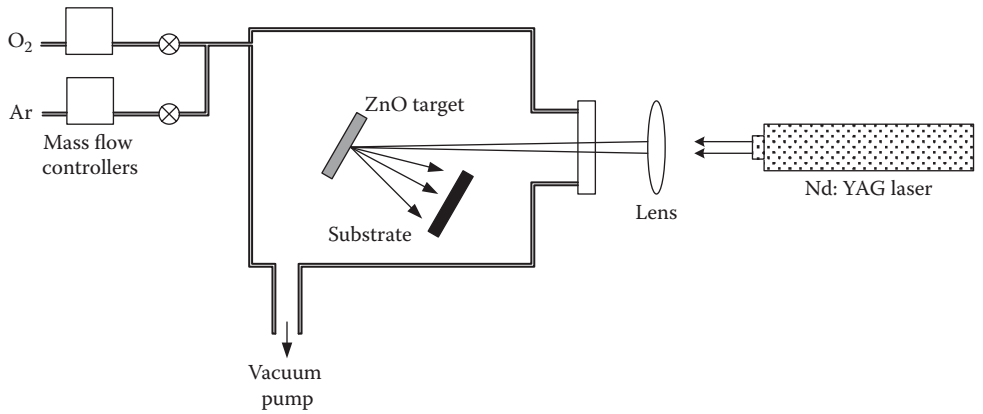


FIGURE 3.12
PLD system for the growth of ZnO films.

3.9 Conclusion

Molecular beam epitaxial and VPE techniques, especially MOVPE, have enabled the realization of a wide range of heteroepitaxial devices and structures. Both afford tremendous flexibility and the ability to deposit thin layers and complex multilayered structures with precise control and excellent uniformity. This enables the practical realization of advanced device structures such as heterojunction, quantum well, and quantum dot devices.

Many heteroepitaxial material combinations with diverse characteristics have been investigated. However, certain aspects of heteroepitaxy appear repeatedly among them, and will be covered in detail in the following chapters. These aspects include nucleation, growth modes, lattice mismatch and strain relaxation, crystal defects, thermal strain, and cracking.

PROBLEMS

1. Calculate the Reynold's number for an MOVPE reactor operating at 0.1 atm and 650°C with 10 slm of H_2 carrier gas, if the reaction chamber is a round tube with a diameter of 10 cm. Determine if laminar or turbulent flow conditions prevail in the reactor.
2. Consider the MOVPE growth of GaAs using TMGa + AsH_3 at 650°C in a round reactor tube with a diameter of 10 cm. Ten stand liters per minute of H_2 carrier gas is used. The mole fraction of TMGa in the reactor is 10^{-4} and the total pressure is 1 atm. (a) Estimate the boundary layer thickness at a distance 1 cm down the susceptor. (b) Estimate the growth rate. Assume the diffusivity of TMGa in hydrogen is 0.31 $cm^2\ s^{-1}$.
3. Repeat Problem 2 for the case of $P = 0.1$ atm. Assume that the diffusivity scales as 1/P.
4. Suppose 20 sccm of H_2 is bubbled through a DMZn bubbler maintained at –10°C. (a) Assuming the bubbles become saturated with the vapor of DMZn, estimate the flow of DMZn to the reactor. (b) Calculate the mole fraction of DMZn in the reactor, if the total flow of H_2 carrier gas is 5 slm.
5. Consider MBE growth of GaAs. The Ga effusion cell has a diameter of 2 cm, is located 25 cm from the substrate, and is held at a temperature of 1000°C. (a) Calculate the impingement rate for Ga at the substrate, in atoms per square centimeter per second. (b) Estimate the growth rate. Assume the vapor pressure of Ga at 1000°C is 4×10^{-3} torr.

References

1. S. Whitaker, *Introduction to Fluid Mechanics* (Englewood Cliffs, NJ: Prentice-Hall, 1968).
2. H. M. Manasevit, Single-crystal gallium arsenide on insulating substrates, *Appl. Phys. Lett.*, 12, 156 (1968).
3. H. M. Manasevit and W. I. Simpson, The use of metal-organics in the preparation of semiconductor materials. I. Epitaxial gallium-V compounds, *J. Electrochem. Soc.*, 116, 1725 (1969).

4. H. M. Manasevit, The use of metalorganics in the preparation of semiconductor materials: Growth on insulating substrates, *J. Cryst. Growth*, 13–14, 306 (1972).
5. M. A. Herman and H. Sitter, *Molecular Beam Epitaxy: Fundamentals and Current Status* (New York: Springer-Verlag, 1988).
6. J. H. Jeans, *An Introduction to the Kinetic Theory of Gases* (Cambridge: University Press, 1967).
7. B. B. Dayton, Gas flow patterns at entrance and exit of cylindrical tubes, in *1956 National Symposium on Vacuum Technology Transactions*, eds. E. S. Perry and T. H. Devant (Oxford: Pergamon Press, 1957), p. 5.
8. T. U. M. S. Murthy, N. Miyamoto, M. Shimbo, and J. Nishizawa, Gas-phase nucleation during the thermal decomposition of silane in hydrogen, *J. Cryst. Growth*, 33, 1 (1976).
9. R. P. Ruth, J. C. Marinace, and W. C. Dunlap Jr., Vapor-deposited single-crystal germanium, *J. Appl. Phys.*, 31, 995 (1960).
10. E. F. Cave and B. R. Czorny, Epitaxial deposition of silicon and germanium layers by chloride reduction, *RCA Rev.*, 523 (1963).
11. K. J. Miller and M. J. Grieco, Epitaxial P-type germanium films by the hydrogen reduction of $GeBr_4$, $SiBr_4$, and BBr_3, *J. Electrochem. Soc.*, 110, 1252 (1963).
12. J. M. Hartmann, Y. Bogumilowicz, F. Andrieu, P. Holliger, G. Rolland, and T. Billon, Reduced pressure-chemical vapor deposition of high Ge content $Si_{1-x}Ge_x$ and high C content $Si_{1-y}C_y$ layers for advanced metal oxide semiconductor transistors, *J. Cryst. Growth*, 277, 114 (2005).
13. S. S. Iyer, K. Eberl, A. R. Powell, and B. A. Ek, SiCGe ternary Alloys-Extending Si-based heterostructures, *Microelectron. Eng.*, 19, 351 (1992).
14. C. Li, S. John, E. Quinones, and S. Banerjee, Cold-wall ultrahigh vacuum chemical vapor deposition of doped and undoped Si and $Si_{1-x}Ge_x$ epitaxial films using SiH_2 and Si_2H_6, *J. Vac. Sci. Technol. A*, 14, 170 (1996).
15. Yu. B. Bolkhovityanov, A. S. Deryabin, A. K. Gutakovskii, M. A. Revenko, and L. V. Sokolov, Heterostructures Ge_xSi_{1-x}/Si (001) (x=0.18–0.62) grown by molecular beam epitaxy at a low (350°C) temperature: Specific features of plastic relaxation, *Thin Solid Films*, 466, 69 (2004).
16. J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy, *J. Vac. Sci. Technol. A*, 2, 436 (1984).
17. J. A. Powell, L. G. Matus, and M. A. Kuczmarski, Growth and characterization of cubic SiC single-crystal films on Si, *J. Electrochem. Soc.*, 134, 1558 (1987).
18. S. Nishino and J. Saraie, Heteroepitaxial growth of cubic SiC on a Si substrate using the Si_2H_6-C_2H_2-H_2 system, in *Amorphous and Crystalline Silicon Carbide*, eds. G. L. Harris and C. Y.-W. Yang (Berlin: Springer, 1989), p. 45.
19. W. Muench, W. Kurzinger, and I. Pfaffender, Epitaxial deposition of silicon carbide from silicon tetrachloride and hexane, *Thin Solid Films*, 31, 39 (1976).
20. P. Liaw and R. F. Davis, Epitaxial growth and characterization of β-SiC thin films, *J. Electrochem. Soc.*, 132, 642 (1985).
21. K. Ikoma, M. Yamanaka, H. Yamaguchi, and Y. Shichi, Heteroepitaxial growth of β-SiC on Si(111) by CVD using a CH_3Cl-SiH_4-H_2 gas system, *J. Electrochem. Soc.*, 138, 3031 (1991).
22. P. Rai-Choudhury and N. P. Formigoni, β-Silicon carbide film, *J. Electrochem. Soc.*, 116, 1440 (1969).
23. S. Nishino and J. Saraie, Heteroepitaxial growth of cubic SiC on a Si substrate using methyltrichlorosilane, in *Springer Proceedings in Physics*, eds. M. M. Rahman, C. Y. Yang, and G. L. Harris, Vol. 43 (Berlin: Springer, 1989), p. 8.
24. T. Chassagne, G. Ferro, D. Chaussnde, F. Cauwet, Y. Monteil, and J. Bouix, A comprehensive study of SiC growth processes in a VPE reactor, *Thin Solid Films*, 402, 83 (2002).
25. R. S. Kern, S. Tanaka, L. B. Rowland, and R. F. Davis, Reaction kinetics of silicon carbide deposition by gas-source molecular-beam epitaxy, *J. Cryst. Growth*, 183, 581 (1998).
26. R. F. Davis, S. Tanaka, L. B. Rowland, R. S. Kern, Z. Sitar, S. K. Ailey, and C. Wang, Growth of SiC and III-V nitride thin films via gas-source molecular beam epitaxy and their characterization, *J. Cryst. Growth*, 164, 132 (1996).

27. R. S. Kern, S. Tanaka, L. B. Rowland, and R. F. Davis, Reaction kinetics of silicon carbide deposition by gas-source molecular-beam epitaxy, *J. Cryst. Growth*, 183, 581 (1998).
28. N. Kuroda, K. Shibahara, W. S. Yoo, S. Nishino, and H. Matsunami, *Extended Abstracts of the 34th Spring Meeting of Japan Society Applied Physics and Related Societies*, Tokyo (1987), p. 135 (in Japanese).
29. N. Kuroda, K. Shibahara, W. S. Yoo, S. Nishino, and H. Matsunami, *Extended Abstracts of the 19th Conference on Solid State Devices and Materials*, Tokyo (1987), p. 227.
30. C. P. Kuo, R. M. Cohen, K. L. Fry, and G. B. Stringfellow, OMVPE growth of GaInAs, *J. Cryst. Growth*, 64, 461 (1983).
31. C. P. Kuo, J. S. Yuan, R. M. Cohen, J. Dunn, and G. B. Stringfellow, Organometallic vapor phase epitaxial growth of high purity GaInAs using trimethylindium, *Appl. Phys. Lett.*, 44, 550 (1984).
32. K. Yuan, K. Radhakrishnan, H. Q. Zheng, and G. I. Ng, Metamorphic $In_{0.5}Al_{0.5}As/In_{0.53}Ga_{0.47}As$ high electron mobility transistors on GaAs with $In_xGa_{1-x}P$ graded buffer, *J. Vac. Sci. Technol. B*, 19, 2119 (2001).
33. S. Kondo, S.-I. Matsumoto, and H. Nagai, 660 nm $In_{0.5}Ga_{0.5}P$ light-emitting diodes on Si substrates, *Appl. Phys. Lett.*, 53, 279 (1988).
34. B. R. Bennett, R. Magno, J. B. Boos, W. Kruppa, and M. G. Ancona, Antimonide-based compound semiconductors for electronic devices: A review, *Solid State Electron.*, 49, 1875 (2005).
35. L. M. Fraas, R. Ballantyne, J. Samaras, and M. Seal, Electric power production using new GaSb photovoltaic cells with extended infrared response, *AIP Conf. Proc.*, 321, 44 (1994).
36. Y. Paltiel, A. Sher, A. Raizman, S. Shusterman, M. Katz, A. Zemel, Z. Calahorra, and M. Yassen, Metalorganic vapor phase epitaxy InSb p^+nn^+ photodiodes with low dark current, *Appl. Phys. Lett.*, 84, 5419 (2004).
37. T. D. McLean, T. M. Kerr, D. I. Westwood, J. D. Grange, and I. J. Murgatroyd, *Inst. Phys. Conf. Ser.*, 74, 145 (1984).
38. H. Ehsani, I. Bhat, C. Hitchcock, J. Borrego, and R. Gutmann, Characteristics of GaSb and GaInSb layers grown by metalorganic vapor phase epitaxy, *AIP Conf. Proc.*, 358, 423 (1996).
39. H. Ehsani, I. Bhat, R. Gutmann, and G. Charache, p-Type GaSb and $Ga_{0.8}In_{0.2}Sb$ layers grown by metalorganic vapor phase epitaxy using silane as the dopant source, *Appl. Phys. Lett.*, 69, 3863 (1996).
40. Y. Koide, H. Itoh, N. Sawaki, I. Akasaki, and M. Hashimoto, Epitaxial growth and properties of $Al_xGa_{1-x}N$ by MOVPE, *J. Electrochem. Soc.*, 133, 1956 (1986).
41. Y. Dikme, A. Szymakowski, H. Kalisch, E. V. Lutsenko, V. N. Zubialevich, G. P. Yablonskii, H. M. Chern, C. Schaefer, R. Jansen, and M. Heuken, Investigation of GaN on Si(111) for optoelectronic applications, *Proc. SPIE*, 4996, 57 (2003).
42. H. Amano, N. Sawaki, I. Akasaki, and Y. Toyoda, Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer, *Appl. Phys. Lett.*, 48, 353 (1986).
43. H. Amano, I. Akasaki, K. Hiramatsu, and Sawaki, Effects of the buffer layer in metalorganic vapour phase epitaxy of GaN on sapphire substrate, *Thin Solid Films*, 163, 415 (1988).
44. Y. Koide, N. Itoh, X. Itoh, N. Sawaki, and I. Akasaki, Effect of AlN buffer layer on AlGaN/α–Al_2O_3 heteroepitaxial growth by metalorganic vapor phase epitaxy, *Jpn. J. Appl. Phys.*, 27, 1156 (1988).
45. S. Nakamura, GaN growth using GaN buffer layer, *Jpn. J. Appl. Phys.*, 30, L1705 (1991).
46. N. Kuznia, M. A, Khan, D. T. Olsen, R. Kaplan, and J. Freitas, Influence of buffer layers on the deposition of high quality single crystal GaN over sapphire substrates, *J. Appl. Phys.*, 73, 4700 (1993).
47. S. Keller, B. P. Keller, Y.-F. Wu, B. Heying, D. Kapolnek, J. S. Speck, U. K. Mishra, and S. P. DenBaars, Influence of sapphire nitridation on properties of gallium nitride grown by metalorganic chemical vapor deposition, *Appl. Phys. Lett.*, 68, 1525 (1996).
48. Y. V. Shvyd'ko, M. Lucht, E. Gerdau, M. Lerche, E. E. Alp, W. Sturhahn, J. Sutter, and T. S. Toellner, Measuring wavelengths and lattice constants with the Mössbauer wavelength standard, *J. Synchrotron Rad.*, 9, 17 (2002).

49. I. Akasaki, H. Amano, Y. Koide, K. Hiramatsu, and N. Sawaki, Effects of AlN buer layer on crystallographic structure and on electrical and optical properties of GaN and $Ga_{1-x}Al_xN$ ($0 < x \leq 0.4$) films grown on sapphire substrate by MOVPE, *J. Cryst. Growth*, 209 (1989).
50. S. Nakamura, Y. Harada, and M. Seno, Novel metalorganic chemical vapor deposition system for GaN growth, *Appl. Phys. Lett.*, 58, 2021 (1991).
51. C.-R. Lee, S.-J. Son, I.-H. Lee, J.-Y. Leem, and S. K. Noh, High-quality GaN epilayers grown by newly designed horizontal counter-flow MOCVD reactor, *J. Cryst. Growth*, 182, 11 (1997).
52. K. Nishida, S. Haneda, K. Hara, H. Munekata, and H. Kukimoto, MOVPE of GaN using a specially designed two-flow horizontal reactor, *J. Cryst. Growth*, 170, 312 (1997).
53. X. Wang and A. Yoshikawa, Molecular beam epitaxy growth of GaN, AlN, and InN, *Prog. Cryst. Growth Char. Mater.*, 48/49, 42 (2004).
54. J. T. Torvik, M. Leksono, J. I. Pankove, B. V. Zeghbroeck, H. M. Ng, and T. D. Moustakas, Electrical characterization of GaN/SiC n-p heterojunction diodes, *Appl. Phys. Lett.*, 72, 1371 (1998).
55. N. Gogneau, E. Sarigiannidou, E. Monroy, S. Monnoye, H. Mank, and B. Daudin, Surfactant effect of gallium during the growth of GaN on AlN (0001) by plasma-assisted molecular beam epitaxy, *Appl. Phys. Lett.*, 85, 1421 (2004).
56. E. Monroy, N. Gogneau, F. Enjalbert, F. Fossard, D. Jalabert, E. Bellet-Amalric, L. S. Dang, and B. Daudin, Molecular-beam epitaxial growth and characterization of quaternary III-nitride compounds, *J. Appl. Phys.*, 94, 3121 (2003).
57. H. M. Ng, D. Doppalapudi, D. Korakakis, R. Singh, and T. D. Moustakas, MBE growth and doping of III-V nitrides, *J. Cryst. Growth*, 189, 349 (1998).
58. G. Namkoong, W. A. Doolittle, and A. S. Brown, Incorporation of Mg in GaN grown by plasma-assisted molecular beam epitaxy, *Appl. Phys. Lett.*, 77, 4386 (2000).
59. J. S. Harris Jr., The opportunities, successes and challenges for GaInNAsSb, *J. Cryst. Growth*, 278, 3 (2005).
60. W. Zheng, Y. Liao, L. Li, Q. Yu, G. Wang, Y. Li, and Z. Fu, Structure and properties of ZnO films grown on Si substrates with low temperature buffer layers, *Appl. Surf. Sci.*, 253, 2765 (2006).
61. H. J. Ko, Y. F. Chen, Z. Zhu, T. Hanada, and T. Yao, Effects of a low-temperature buffer layer on structural properties of ZnO epilayers grown on (111) CaF_2 by two-step MBE, *J. Cryst. Growth*, 208, 389 (2000).
62. H. F. Liu and S. J. Chua, Efects of low-temperature-bufer, RF-power, and annealing on structural and optical properties of $ZnO/Al_2O_3(0001)$ thin films grown by RF-magnetron sputtering, *J. Appl. Phys.*, 106, 023511 (2009).
63. W. E. Hoke, P. J. Lemonias, and R. Taczewski, Metalorganic growth of high-purity HgCdTe films, *Appl. Phys. Lett.*, 45, 1092 (1984).
64. I., B. Bhat, N. R. Taskar, and S. K. Ghandhi, The organometallic heteroepitaxy of CdTe and HgCdTe on GaAs substrates, *J. Vac. Sci. Technol. A*, 4, 2230 (1986).
65. J. P. Faurie, S. Sivanathan, M. Boukerche, and J. Reno, Molecular beam epitaxial growth of high quality HgTe and $Hg_{1-x}Cd_xTe$ onto GaAs (001) substrates, *Appl. Phys. Lett.*, 45, 1307 (1984).
66. K. Nishitani, R. Okhata, and T. Murotani, Molecular beam epitaxy of CdTe and $Hg_{1-x}Cd_xTe$ on GaAs (100), *J. Electron. Mater.*, 12, 619 (1983).
67. T. H. Myers, Y. Lo, R. N. Bicknell, and J. F. Schetzina, Growth of CdTe films on sapphire by molecular beam epitaxy, *Appl. Phys. Lett.*, 42, 247 (1983).
68. J. T. Cheung and T. J. Magee, Recent progress on LADA growth of HgCdTe and CdTe epitaxial layers, *J. Vac. Sci. Technol. A*, 1, 1604 (1983).
69. H. A. Mar, K. T. Chee, and N. Salansky, CdTe films on (001) GaAs:Cr by molecular beam epitaxy, *Appl. Phys. Lett.*, 44, 237 (1984).
70. S. K. Ghandhi, I. B. Bhat, and N. R. Taskar, Growth and properties of $Hg_{1-x}Cd_xTe$ on GaAs substrates by organometallic vapor-phase epitaxy, *J. Appl. Phys.*, 59, 2253 (1986).
71. K. Yasuda, H. Hatano, T. Ferid, M. Minamide, T. Maejima, and K. Kawamoto, Growth characteristics of (100) HgCdTe layers in low-temperature MOVPE with ditertiarybutyltelluride, *J. Cryst. Growth*, 166, 612 (1996).

72. V. S. Varavin, S. A. Dvoretsky, V. I. Liberman, N. N. Mikhailov, and Yu. G. Sidorov, Molecular beam epitaxy of high quality $Hg_{1-x}Cd_xTe$ films with control of the composition distribution, *J. Cryst. Growth*, 159, 1161 (1996).
73. K. Hirano, M. Fujita, M. Sasajima, T. Kosaka, and Y. Horikoshi, ZnO epitaxial films grown by flux-modulated RF-MBE, *J. Cryst. Growth*, 301, 370 (2007).
74. J. S. Park, S. K. Hong, T. Minegishi, I. H. Im, S. H. Park, T. Hanada, J. H. Chang, M. W. Cho, and T. Yao, The high quality ZnO growth on c-Al_2O_3 substrate with Cr_2O_3 buffer layer using plasma-assisted molecular beam epitaxy, *Appl. Surf. Sci.*, 254, 7786 (2008).
75. J. S. Park, S. K. Hong, I. H. Im, J. S. Ha, H. J. Lee, S. H. Park, J. H. Chang, M. W. Cho, and T. Yao, Growth of high-quality ZnO films on Al_2O_3 (0001) by plasma-assisted molecular beam epitaxy, *J. Cryst. Growth*, 311, 2163 (2009).
76. F. Xiu, Z. Yang, D. Zhao, J. Liu, K. A. Alim, A. A. Balandin, M. E. Itkis, and R. C. Haddon, ZnO growth on Si with low-temperature ZnO buffer layers by ECR-assisted MBE, *J. Cryst. Growth*, 286, 61 (2006).
77. D. C. Look, D. C. Reynolds, C. W. Litton, R. L. Jones, D. B. Eason, and G. Cantwell, Characterization of homoepitaxial p-type ZnO grown by molecular beam epitaxy, *Appl. Phys. Lett.*, 81, 1830 (2002).
78. J. R. Botha, K. T. Roro, C. Weichsel, A. W. R. Leitch, and J. Weber, Arsenic-related recombination in MOVPE-grown ZnO/GaAs films, *Superlat. Microstruct.*, 42, 26 (2007).
79. P. Kuznetsov, V. Lusanov, G. Yakushcheva, V. Jitov, L. Zakharov, I. Kotelyanskii, and V. Kozlovsky, MOVPE growth and study of ZnO, ZnMgO epilayers and ZnO/ZnMgO MQW structures, *Phys. Status Solidi C*, 7, 1568 (2010).
80. K. Ogata, K. Maejima, Sz. Fujita, and Sg. Fujita, ZnO growth toward optical devices by MOVPE using N_2O, *J. Electron. Mater.*, 30, 659 (2001).
81. K. Ogata, S.-W. Kim, Sz. Fujita, and Sg. Fujita, ZnO growth on Si substrates by metalorganic vapor phase epitaxy, *J. Cryst. Growth*, 240, 112 (2002).
82. H. Yamaguchi, T. Shitara, T. Komiyama, Y. Chonan, and T. Aoyama, ZnO films deposited by optimized PLD technique with bias voltages, *Phys. Status Solidi C*, 7, 326 (2010).
83. J.-H. Leem, D.-H. Lee, and S. Y. Lee, Properties of N-doped ZnO grown by DBD-PLD, *Thin Solid Films*, 518, 1238 (2009).
84. Y.-S. Jung, O. Kononenko, J.-S. Kim, and W.-K. Choi, Two-dimensional growth of ZnO epitaxial films on c-Al_2O_3 (0001) substrates with optimized growth temperature and low-temperature buffer layer by plasma-assisted molecular beam epitaxy, *J. Cryst. Growth*, 274, 418 (2005).
85. H. P. Maruska and J. J. Tietjen, The preparation and properties of vapor-deposited single-crystalline GaN, *Appl. Phys. Lett.*, 15, 327 (1969).

4

Surface and Chemical Considerations in Heteroepitaxy

4.1 Introduction

Heteroepitaxy differs from homoepitaxy in that it requires the nucleation of a new phase A on a foreign substrate B. Because of this, the surface chemistry and physics play important roles in determining the properties of heteroepitaxial deposits, including structural and electrical characteristics, defect densities and structure, and the layer morphology.

At the typical temperatures employed for epitaxial growth, the substrate surface may undergo a reconstruction, which is an atomic-scale change in surface structure. The structure of the reconstructed surface may depend on the temperature and absorbed species. Moreover, in some cases it has been found that the initial structure of the reconstructed substrate affects the structural and electrical properties of thick heteroepitaxial films grown on its surface.

The importance of the surface and its structure in heteroepitaxial growth has been highlighted by recent studies of the interfacial misfit (IMF) dislocation array mode, which has been observed in the heteroepitaxy of GaSb on GaAs(001).[1] In the case of the IMF growth mode, lattice relaxation is effected by regular arrays of 90° dislocations along the [110] and [$1\bar{1}0$] directions with a spacing of 5.6 nm, corresponding to 13 lattice sites of GaSb and 14 lattice sites of GaAs, resulting in a very low threading dislocation density. The mechanism of IMF growth, though incompletely understood at the present time, has been proposed to relate to the self-assembly of Sb adatoms on the surface, and lending support to this theory is the observation that the optimal IMF array is formed when Sb interacts with a (2×8)-reconstructed Ga-rich GaAs(001) surface.[1]

On vicinal (tilted) substrates, the structure of steps and kinks on the surface can have an important influence on the heteroepitaxial growth. For example, in the growth of a polar semiconductor on a nonpolar substrate, inversion domain boundaries (also referred to as antiphase domain boundaries) may develop due to the lower symmetry of the heteroepitaxial crystal. However, it has been found that this behavior can be controlled by the proper choice of the substrate tilt and direction.

Either energetics or kinetics may control the nucleation and growth mode for heteroepitaxy. Traditionally, three possible growth modes have been identified as Frank–van der Merwe (FM), Volmer–Weber (VW), and Stranski–Krastanov (SK) growth. Recently, another mode of growth ("ripening") has been identified in which large islands grow, possibly in the presence of a wetting layer or smaller, stable islands. The FM mode involves layer-by-layer growth, giving smooth interfaces; it is desirable for most device applications and is mandatory for quantum well layers. VW (island growth) and SK (islanding on a

continuous wetting layer) growth are undesirable for most applications; however, they can be useful for the fabrication of quantum dot devices. In all of these situations, however, understanding the growth mode is important for the design of the device fabrication process.

Another important aspect of surface chemistry in heteroepitaxy involves the use of surfactants. These are atomic (or molecular) species that are preferentially segregated to the surface instead of being incorporated into the growing crystal. Such an adsorbed species may have a dramatic effect on the growth even after modest exposure of the surface (~1 monolayer). If the incorporation and evaporation of the surfactant are both negligible, then even a short exposure to the surfactant species can alter the surface chemistry throughout the growth of a thick film, affecting the composition, structure, and morphology of the resulting material.

This chapter reviews some of the important aspects of surfaces and their chemical considerations, as they apply to heteroepitaxy. Surface reconstructions will be discussed, including some general principles, but also the specifics of some important surfaces for semiconductor epitaxy. Inversion domain boundaries are reviewed; it is also shown that, at least to some extent, the surface structure controls the introduction of inversion domains in a growing polar semiconductor on a nonpolar substrate. Nucleation theory is reviewed, starting with the classical model based on the macroscopic surface and interface energies. Then, the atomistic model for nucleation is summarized as it applies to heteroepitaxy. Growth modes are closely related to nucleation, but are covered in a separate section. The discussion of growth modes includes a thermodynamic (equilibrium) treatment, which allows the calculation of a growth mode phase diagram for heteroepitaxy. This equilibrium treatment is also extended to include misfit dislocations, with application to SiGe. Then, a kinetic model is presented, which describes the conditions for the onset of island growth on top of a wetting layer. In the next section, surfactants are discussed, starting with some general principles and ending with important applications to heteroepitaxy. In the final section of the chapter, several aspects of quantum dot fabrication are addressed, including self-assembly, self-organization, and precision placement.

4.2 Surface Reconstructions

The surface of a semiconductor substrate will generally take on a structure different from that of a truncated bulk crystal. The driving force for this is energy minimization. In some cases, the rearrangement is rather subtle, altering neither the periodicity nor the symmetry of the surface. This is referred to as surface relaxation. In other cases, the rearrangement is such that it changes the periodicity, and perhaps the symmetry, of the surface, and is called surface reconstruction. Often, such reconstructions have been attributed to dimerization of the surface atoms, which serves to reduce the number of unsaturated dangling bonds. Reconstruction can be readily detected by electron diffraction techniques, including low-energy electron diffraction (LEED) and reflection high-energy electron diffraction (RHEED).

Just as three-dimensional (3-D) crystal structures belong to 1 of the 14 Bravais lattices, surface atomic arrangements can be considered to belong to 1 of 5 types of surface nets. These are the square, rectangular, centered rectangular, hexagonal, and oblique nets, and their unit meshes are illustrated in Figure 4.1.

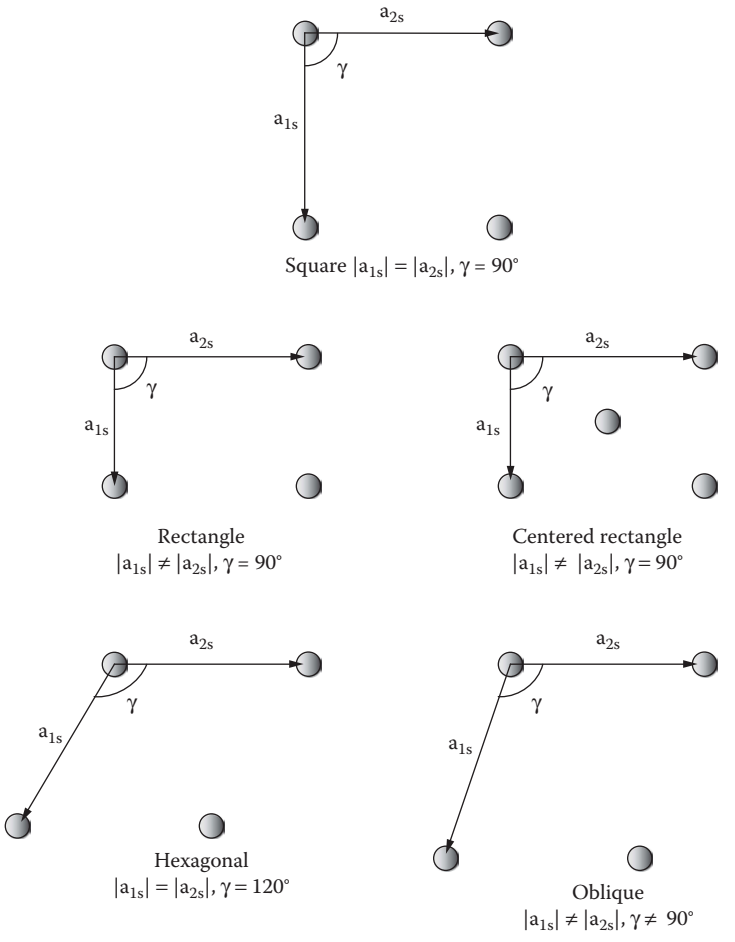


FIGURE 4.1
Unit cells of the five surface nets. (Adapted from E. A. Wood, *J. Appl. Phys.*, 35, 1306 [1964]. With permission. Copyright 1964, American Institute of Physics.)

The assignment of the surface structure involves the identification of its symmetry (the shape and dimensions of the unit cell) and the determination of the positions of the atoms within the unit cell. The former is readily found from the positions of diffracted electron beams, but the latter requires an analysis of the diffracted intensities. This interpretation is not straightforward, because phase information is lost in the diffraction pattern. Thus, it is necessary to guess the structure, simulate the diffracted intensities, and then compare the calculated and measured results. Several iterations may be required, but even then it may not be possible to identify a unique solution. In most cases, the detailed atomic structures of reconstructed semiconductor surfaces remain unknown. However, recent advances in lateral force microscopy (LFM)[2,3] (a variation of atomic force microscopy) have made it possible to observe reconstructed surfaces with true atomic resolution.[4] By these measurements, it is possible to elucidate the detailed structures of the unit meshes in reconstructed surfaces. Routine electron diffraction measurements can then be used to distinguish between the types of surface structures without a need for such detailed analysis.

4.2.1 Wood's Notation for Reconstructed Surfaces

Usually, reconstructed surfaces are classified using Wood's notation.[5] Suppose a_S is the surface mesh with unit translations a_{1S} and a_{2S}. Further, suppose that the mesh of an unreconstructed surface ("bulk exposed plane") is a_B with unit translations a_{1B} and a_{2B}. In Wood's notation, the relationship between the reconstructed mesh and the mesh of the bulk exposed plane is expressed as $(a_{1S}/a_{1B} \times a_{2S}/a_{2B})R$, where the R indicates a rotation of the surface mesh with respect to the bulk and is followed by the value of this rotation in degrees. (If there is no rotation of the surface mesh, R is omitted.) The notation $c(a_{1S}/a_{1B} \times a_{2S}/a_{2B})R$ is used to denote a centered mesh. Wood's notation is applicable to clean surfaces, but can also be extended to situations involving an absorbate.

Two example meshes are shown in Figure 4.2, along with their Wood's classification. Figure 4.2a shows adsorbed oxygen on a Ni(110) surface, with a centered (2×2) mesh and no rotation. The notation for this surface structure is Ni(110)$c(2 \times 2)$–O. Figure 4.2b shows adsorbed oxygen on a Pt(100) surface, with a $(\sqrt{2} \times 2\sqrt{2})$ mesh, rotated by 45°. The notation in this case is Pt(100)$(\sqrt{2} \times 2\sqrt{2})R45°$–O.

In general, the structure of a reconstructed surface will depend on the temperature as well as the beam fluxes (molecular beam epitaxy [MBE]) or source flows (metaloganic vapor phase epitaxy [MOVPE]). Different reconstructions may be observed at different temperatures during the pregrowth processing in an epitaxial reactor. Also, kinetic limitations

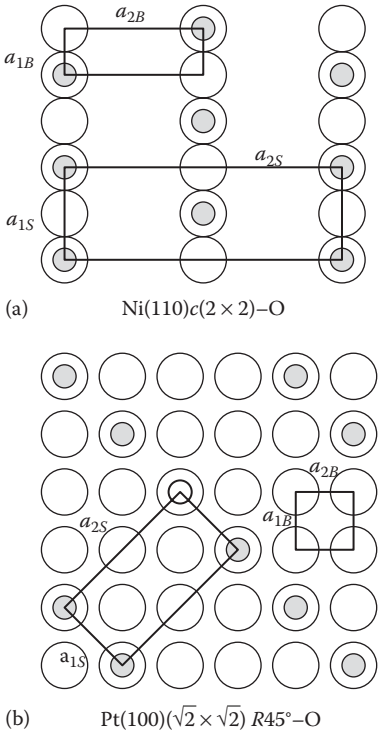


FIGURE 4.2
Surface structures for the illustration of Wood's notation. (a) Ni(110) surface with adsorbed oxygen having a centered (2×2) mesh and no rotation, denoted Ni(110)$c(2 \times 2)$–O. (b) Pt(100) surface with adsorbed oxygen having a $(\sqrt{2} \times 2\sqrt{2})$ mesh, rotated by 45°, denoted Pt(100)$(\sqrt{2} \times 2\sqrt{2})R45°$–O. (Adapted from M. Prutton, *Introduction to Surface Physics* [Oxford: Oxford University Press, 1994]. Used with permission of the Oxford University Press.)

may prevent the surface from taking on the equilibrium structure. Thus, a reconstructed structure that is stable only at high temperature may be "frozen in" at a lower temperature.

The presence of absorbed species will generally affect the surface structure. Many surface structures that have been observed on Si only exist in the presence of adsorbed impurities. These could have important implications for nucleation and heteroepitaxy. One application area is surfactant-mediated epitaxy (SME). For example, in the surfactant-mediated heteroepitaxy of Ge on Si(111) using Sb as a surfactant, the adsorption of Sb changes the surface structure[6] from Si(111)(7 × 7) to Si(111)($\sqrt{3} \times \sqrt{3}$)*R*30°–Sb. In general, adsorption of species can alter the surface structure during epitaxy, even in the case of homoepitaxy. For example, in the case of ultra-high-vacuum (UHV) VPE of silicon using silane on Si(111), the surface is found to convert from the Si(111)(7 × 7) reconstruction to a partially hydrogenated Si(111)(1 × 1) structure[7] upon exposure to silane. In this case, therefore, the nucleation and growth occur on different surfaces.

It is clear that the surface structure will influence heteroepitaxial growth in a variety of ways, many of which are poorly understood. It is known, for example, that in VPE growth the growth rate is determined by the concentration of adsorbed species on reactive sites, and this is controlled by the surface atomic structure. Further study in this area is likely to result in progress in the areas of SME, selective epitaxy, and quantum devices.

4.2.2 Experimental Observations

Some of the commonly used substrate surfaces have been studied extensively; some of the results of these investigations are summarized in the following subsections.

4.2.2.1 Si(001) Surface

The clean Si(001) surface is found to assume a (2 × 1)-type reconstruction[8] if heated in an ultrahigh vacuum or a hydrogen ambient. In all probability, the atomic structure of the reconstructed surface is consistent with the pairing model;[9–13] other models have been proposed, however, including the vacancy model.[14] However, the *c*(4 × 2) reconstruction[15] and other structures have also been observed.[16]

The nominal (001) surface is populated with steps, the types and densities of which are determined by the miscut direction and angle. In the absence of an intentional miscut angle, or if the miscut angle is small (<2°), monatomic (single-layer) steps exist on the surface. These single-layer steps separate the S_A and S_B domains, which exhibit (2 × 1) and (1 × 2) reconstructions, respectively. Here, the surface meshes have the same structure and symmetry but are rotated with respect to each other by 90°. Vicinal Si(001) substrates with miscut angles of 2°–11° are found to have predominantly double-layer steps if annealed at 600°C–1200°C.[17–20] As a consequence, these surfaces have a single type of surface domain with the (2 × 1) reconstruction. Equilibrium calculations have been used to determine the temperature versus miscut angle phase diagram for the Si(001) surface.[21] These results predict that the double-step phase is more stable than the single-step phase at typical epitaxial growth temperatures and with typical miscut angles.

Kinks will generally exist along the steps of a nominal Si(001) surface. For an Si(001) surface with a tilt exactly toward the [110] axis, the equilibrium concentration of such kinks is expected to be small at typical growth temperatures. However, the deliberate choice of a different offcut direction introduces a significant "forced" kink density. This forced kink density increases as the offcut azimuth is moved away from the [110] toward either the [100] or [010].[17]

4.2.2.2 Si(111) Surface

The Si(111) surface is known to exhibit several reconstructions. After cleavage under ultrahigh vacuum at room temperature, an Si(111) (2×1) reconstruction is found.[22] Upon heating to a temperature above 350°C, this surface is found to make an irreversible phase change to a Si(111) (7×7).[23] Detailed studies of this structure by the specular scattering of atomic He reveal that both the first two double layers of atoms are exposed, and that a large fraction of the top layer is missing.[24] Recently, dynamic LFM with atomic resolution has been applied to image the Si(111) (7×7).[25] Such an image is shown in Figure 4.3.

Impurity-stabilized Si(111) surfaces have been found to exhibit a number of reconstructions with periodicities of 1, $\sqrt{3}$, $2\sqrt{3}$, 4, 5, 6, 12, 18, and 24.[26] In the case of nickel stabilization, even a Si(111)$(\sqrt{19} \times \sqrt{19})$ has been observed.[27] Impurity-stabilized surface reconstructions could have important implications for nucleation and heteroepitaxial growth.

4.2.2.3 Ge(111) Surface

As with Si(111), this surface of Ge is observed to take on a (2×1) configuration after cleaving in vacuum. This surface, upon annealing in hydrogen, takes on a Ge(111) (8×8) structure.[28] The (2×8) structure has often been reported based on LEED measurements.[29] However, it has been shown that the observed LEED pattern corresponds to the (8×8) reconstruction, but with certain structure factor cancellations that are due to distortions within the unit mesh. A number of other impurity-induced surface structures have also been observed, as in the case of Si(111).

4.2.2.4 6H-SiC(0001) Surface

6H-SiC(0001) wafers (Si face) are commonly used as substrates for GaN heteroepitaxy, as well as SiC homoepitaxy. The behavior of this surface is quite complex, and many different reconstructions have been observed. The surface structures most commonly found[30] are the $(\sqrt{3} \times \sqrt{3})R30°$ and (3×3) under Si-rich conditions, and the $(6\sqrt{3} \times 6\sqrt{3})R30°$ under C-rich conditions. Other reconstructions have also been reported, including the (3×1),[31] (6×6),[32] and even $(2\sqrt{3} \times 2\sqrt{13})$.[33] These various structures exhibit differences in mobilities for diffusing surface species, and also behave differently with respect to surface-adsorbed surfactants. They are therefore expected to influence both nucleation and epitaxial growth. In the case of GaN grown on 6H-SiC(0001) with an AlN buffer layer, it has been shown that the nature of the 6H-SiC surface reconstruction has a strong effect on the AlN lattice relaxation, and therefore the crystalline quality of the GaN overlayer.[34]

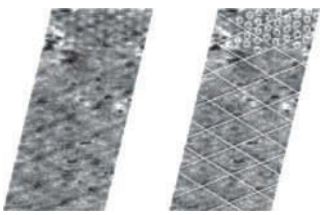


FIGURE 4.3
(a) LFM image of the Si(111) (7×7) surface. (b) The same image overlaid with white circles (individual atoms) and white diamonds (unit cells). (Reprinted from S. Kawai et al., *Appl. Phys. Lett.*, 87, 173105 (2005). With permission. Copyright 2005, American Institute of Physics.)

The tendencies toward these various surface reconstructions may be controlled to a great extent by the methods of surface preparation. For example, Hartman et al.[35] investigated various gas phase treatments and the resulting surface structures. They found that etching at 1500°C–1640°C and 1 atm produced a 1×1 surface. Further, annealing in hydrogen at about 1000°C and 1 atm gave rise to a conversion to the $(\sqrt{3} \times \sqrt{3})R30°$ reconstruction. Subsequent exposure to SiH_4 under UHV conditions was found to convert this surface to a (3×3) structure, but conversion to a (1×1) surface resulted from additional annealing. Suda et al.[34] observed a $(\sqrt{3} \times \sqrt{3})R30°$ structure after etching in HCl/H_2 at 1300°C, but found that a (1×1) surface resulted if the HCl/H_2 etch was followed by a wet HF treatment. Lu et al.[36] observed the $(\sqrt{3} \times \sqrt{3})R30°$ reconstruction after etching in atomic hydrogen at 650°C. Kim et al.[31] investigated the structure of the 6H-SiC(0001) surface after a Ga flash-off process[37] prior to epitaxial growth. The Ga flash-off, which is done in the growth chamber prior to MBE growth, typically consists of exposing the surface to the Ga source flux at a temperature of 650°C, followed by flashing to 800°C in several cycles.[31] This process serves to convert the oxide to the volatile GaO or GaO_2, which can be removed by flashing in the UHV. They observed the (3×1) reconstruction after a Ga flash-off process, even if the oxide removal was incomplete. Often, however, the surface structure exhibits mixed phases rather a single structure.

The detailed atomic configurations of the 6H-SiC(0001) surface reconstructions are incompletely understood. However, recent investigations by scanning tunneling microscopy (STM) have revealed some important features,[38] and structural models have been proposed for the $(\sqrt{3} \times \sqrt{3})R30°$[39], (3×3)[40,41], $(6\sqrt{3} \times 6\sqrt{3})R30°$[42], and $(2\sqrt{3} \times 2\sqrt{3})R30°$[33] reconstructions. An interesting feature of the $(2\sqrt{3} \times 2\sqrt{3})R30°$ surface is the presence of atomic cracks,[38] which have been observed by STM. These cracks might induce island growth or control the formation of boundaries between islands.

4.2.2.5 3C-SiC(001)

The clean (001) surface of zinc blende SiC has been found to take on a (3×2) reconstruction.[43–45] Following the adsorption of hydrogen, this surface retains its (3×2) symmetry but undergoes a transition from semiconductor to metallic character.[46–48] However, this material has been studied to a lesser extent than the hexagonal polytypes.

4.2.2.6 3C-SiC(111)

There are also few reported studies of the (111) surface of zinc blende SiC. Nonetheless, several groups have investigated the surfaces of thin 3C-SiC(111) produced on Si(111) substrates. These layers are either deposited by chemical vapor deposition (CVD) or produced by the reaction of the surface with fullerene molecules such as C_{60} or C_{70}.

Hu et al.[49] prepared a 3C-SiC(111) layer by the thermal reaction of the fullerene molecule C_{60} with the clean surface of Si(111). By using high-resolution electron energy loss spectroscopy (HREELS) and STM, they studied the surface structure and its relationship with the surface chemical composition. In these experiments, the C_{60} source was evaporated onto the Si substrate at room temperature. A short anneal at 250°C served to desorb all but one monolayer of the C_{60} from the surface. As the temperature was raised, there was no further change in the surface up to a temperature of 850°C. At 870°C, however, the C_{60} reacted with the surface to form SiC islands. The SiC surface contained a mixture of (2×2) and (2×3) surface reconstructions, and also a mixture of both the C(111) and Si(111) faces. Finally, annealing at or above 1100°C resulted in the Si-terminated 3C-SiC(111) (3×3) reconstruction.

4.2.2.7 GaN(0001)

Due to the limited availability of single-crystal GaN substrates, most studies of the GaN surface have been conducted with heteroepitaxial layers, usually on sapphire(0001) or 6H-SiC(0001), but occasionally other substrates have been utilized, such as $ScAlMgO_4$.[50] In some studies, it was reported that only the (1 × 1) structure could be observed.[51–53] Other groups have found other reconstructions on the basis of RHEED studies, including (2 × 1), (2 × 2), (2 × 3), (3 × 2), (3 × 3), (4 × 4), (5 × 5), (6 × 4), (6 × 6), *c*(6 × 12).[50,54–60] Smith et al.[60] reported that, in the growth of GaN(0001) on *c*-face sapphire by MBE using a radio frequency (RF) plasma source, the polarity of the surface could be determined by observation of the surface reconstruction. They reported that the (1 × 1), (3 × 3), (6 × 6), and *c*(6 × 12) structures were associated with the N face, whereas the (2 × 2), (5 × 5), and (6 × 4) structures were indicative of the Ga face.

4.2.2.8 Zinc Blende GaN(001)

Zinc blende GaN(001) has been grown heteroepitaxially on several substrates, including Si(001), 3C-SiC(001), and GaAs(001). There have been few reports on the surface structure, but Feuillet et al.[61] found that as-grown MBE GaN(001) on cubic SiC(001) exhibited a (4 × 1) reconstruction but underwent an irreversible transformation to a (2 × 2) structure upon exposure to an As beam.

4.2.2.9 GaAs(001)

GaAs(001) substrates, as prepared for heteroepitaxy, typically exhibit a (2 × 4) reconstruction as observed by the RHEED pattern, which is indicative of an arsenic-stabilized surface.[1,62,63] This surface reconstruction has been studied extensively due to its importance in heteroepitaxy. At an As coverage of 0.75 monolayer (ML), the GaAs(001) and InAs(001) (2 × 4) reconstructions assume the β2 structure.[64–76] The structure contains two As dimers in the first (top) atomic layer and one As dimer in the third layer, per unit cell. At a lower As coverage of 0.5 ML, the α structure[73,77,76] is observed. The α structure contains two As dimers in the first layer and two Ga dimers in the second layer, per unit cell. Both reconstructions are observed under the normal conditions of As stabilization, but the β2(2 × 4) reconstruction is generally observed at lower temperatures, whereas the α(2 × 4) phase is found at higher temperatures. Galitsyn et al.[77] studied the β2(2 × 4)→α(2 × 4) phase transition and determined the critical values of temperature and As overpressure. A GaAs(001) surface with the (2 × 4) structure that is held at typical growth temperatures without an arsenic flux rapidly transitions to the (4 × 2) reconstruction, indicating a Ga-rich surface.[1]

4.2.2.10 InP(001)

The InP(001) surface has been found to exhibit a (2 × 1) reconstruction.[78–80] This structure has been detected on the surface of MOVPE InP(001) by LEED and x-ray photoemission spectroscopy (XPS); infrared spectroscopy of the same sample after deuterium exposure showed that the (2 × 1) structure was hydrogen stabilized.[81] *Ab initio* computations have shown that the adsorption of the surfactant Sb can lead to a number of surface structures on InP(001), including a number of (2 × 3), (2 × 4), (4 × 3), and (4 × 4) variants. A phase diagram has been developed from these calculations and shows that the equilibrium surface reconstruction depends on the chemical potentials (and therefore partial pressures) of both Sb and P over the surface.[82]

4.2.2.11 Sapphire(0001)

Al_2O_3 (sapphire) substrates are of considerable interest due to their use in the heteroepitaxy of GaN and related III-nitrides, and other semiconductors. The sapphire (0001) face is found to exhibit (1 × 2), (2 × 2), (5 × 5), and $(\sqrt{31} \times \sqrt{31})R \pm 9°$ reconstructions. Several studies have related the initial surface structure to the properties of deposited heteroepitaxial layers. Shen et al.[83] reported that for plasma-assisted MBE GaN layers grown on *c*-face sapphire, the highest electron mobilities were obtained when the substrate showed the (1 × 2) pattern.

In summary, there have been extensive modeling and experimental studies of semiconductor surface structures. It is clear that these surface structures can play important roles in determining the properties and usefulness of heteroepitaxial semiconductor layers. However, an understanding of this relationship is only beginning to emerge.

4.2.3 Surface Reconstruction and Heteroepitaxy

Heteroepitaxy usually involves the creation of an interface between semiconductor crystals having different ionicities. As a consequence, the surface structure of the substrate crystal can have an important effect on the properties of the epitaxial crystal. One reason for this is that, depending on the structure of the reconstructed surface, an electrostatic field may develop at the interface, resulting in a rough interface with degraded electrical properties.[84,85] In the case of heteroepitaxy of a polar semiconductor on a nonpolar substrate, an additional challenge arises; this is the possibility of inversion domain boundaries (IDBs). These two issues will be discussed in this section.

4.2.3.1 Inversion Domain Boundaries

The Si(001) surface nominally exhibits a (2 × 1) reconstruction with two types of domains that are rotated by 90° with respect to each other. The silicon atoms on the surface form oriented dimers along the [110] direction for one type of domain (called S_A terraces) and along the $[1\bar{1}0]$ direction for the rotated type of domain (called S_B terraces).[86] These domain terraces are separated by monatomic steps of height $a/4$. The presence of these monatomic steps has been shown to lead to IDBs in the heteroepitaxial growth of polar semiconductors, including GaAs[87] and AlN[88] on Si(001). This behavior can be understood with the aid of Figure 4.4, which shows a Si(001) with a monatomic step, and the atomic configuration of a polar semiconductor grown over the stepped surface. It was assumed that the growth of the polar semiconductor was initiated by exposure of the surface to the anion, so that all Si atoms at the surface bond to As atoms. As a consequence, the polar semiconductor (zinc blende GaAs) contains an IDB associated with the step in the substrate surface.

One approach to eliminating IDBs that has been proposed in heteroepitaxy of a polar semiconductor on Si(001) is the elimination of monatomic surface steps. Double steps are not thought to result in the production of IDBs, according to the model shown in Figure 4.5. Here, it can be seen that the polar semiconductor (zinc blende GaAs) is expected to grow with the same polarity over both types of terraces. However, it has been argued by Pukite and Cohen that this model is an oversimplification and that inversion domain free material may be grown on vicinal silicon containing single steps.[89]

Si(001) surfaces, completely free from monatomic steps, can be prepared by the hydrogen annealing of vicinal substrates. If the Si(001) substrates are offcut by ≥4° toward a {110} direction, and are annealed for several minutes in hydrogen at a temperature of

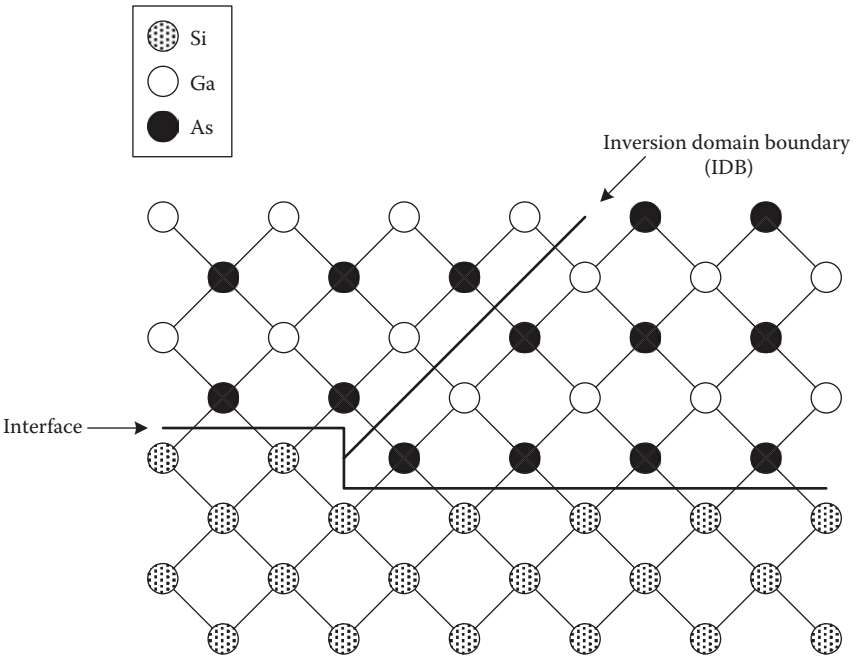


FIGURE 4.4
Growth of a polar semiconductor (zinc blende GaAs) on a Si(001) surface having monatomic steps. IDBs are associated with the monatomic steps.

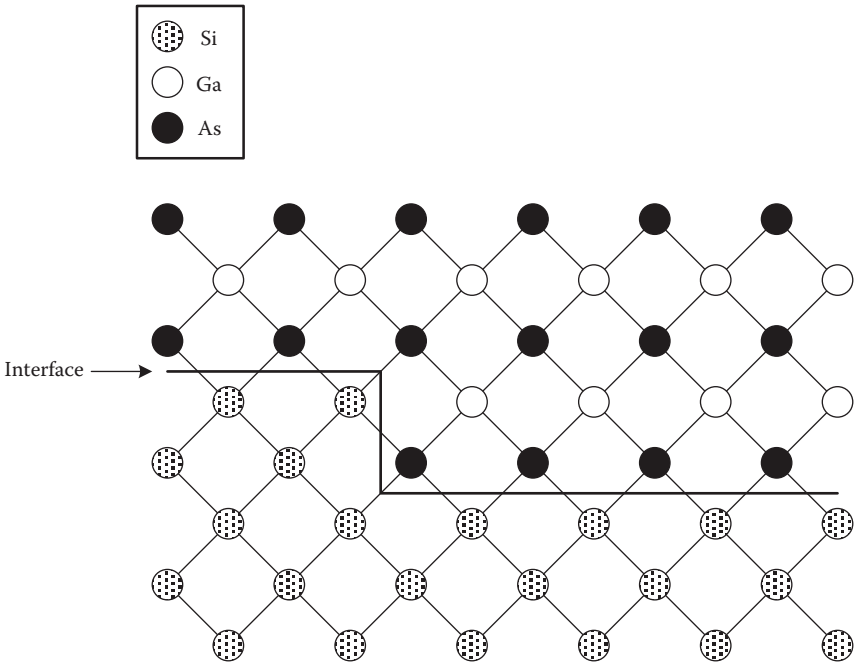


FIGURE 4.5
Growth of a polar semiconductor (zinc blende GaAs) on a Si(001) surface having double steps. The overgrowth is free from IDBs.

900°C–1100°C, only double steps exist on the surface.[20,86,90–92] These surfaces exhibit only a single type of reconstruction terrace, and IDBs can be suppressed by their use, for the growth of CdTe,[93] GaAs, InP,[94] or AlN.[95]

Although the use of vicinal Si(001) substrates, miscut by 4°, has been shown repeatedly to eliminate IDBs in heteroepitaxial layers of III-V semiconductors, a disadvantage of the approach is that it is incompatible with standard complementary metal oxide semiconductor (CMOS) logic circuit fabrication, which uses substrates with the exact (001) orientation.[96] This is an important drawback because an important goal of III-V epitaxy on Si substrates is the integration of III-V optoelectronic components with Si CMOS logic circuits.

Noge et al.[97] found that inversion domain disorder could be eliminated from GaAs grown on Si with "exact" (001) by a pregrowth anneal of the substrate for 30 min at 1000°C, postulating that this caused the elimination of single surface steps by reactions to give rise to double steps.

Another approach is to render the Si surface free from all steps. In some works, it has been shown that the use of annealing at temperatures greater than 1200°C in conjunction with trench etching allows removal of all steps to the edges of exact (001) Si wafers.[98,99] Balakrishnan et al.[96] found that use of trench-edged Si surfaces allowed growth of $Al_yGa_{1-y}Sb$ on exact (001) Si with a significantly reduced density of IDBs even without the use of a high-temperature step. In their demonstration, super-luminal GaSb-based light-emitting diodes (LEDs) were operated with peak spectral output at 2.14 μm and 0.125 mW output power using continuous wave (CW) electrical injection. All the same, the devices exhibited a relatively high series resistance of 27 Ω, which might be attributed to the presence of IDBs at the Si/III–Sb interface.

4.2.3.2 Heteroepitaxy of Polar Semiconductors with Different Ionicities

In the case of heteroepitaxy of a polar semiconductor on another polar semiconductor, IDBs are not expected to occur on the basis of simple energy considerations. Nonetheless, if the two semiconductors have different ionicities, an electric field may develop at the interface, resulting in a rough interface with degraded electrical properties.[84,85]

Farrell et al.[100] studied this phenomenon as it applies to the growth of II-VI materials on GaAs(001) substrates. They considered the case of ZnSe on GaAs for specificity; in this case, the electrons contributed to bonding by each atom will be 2/4, 3/4, 5/4, and 6/4 per bond for Zn, Ga, As, and Se, respectively. A specific case is ZnSe grown with Se initiation on a GaAs(001) surface terminated exclusively with Ga. Here, there will be 1/4 excess electron per bond on the surface. In order to avoid a built-in electric field at the interface, an optimal surface structure would result in a 50:50 interface bond ratio, with a total of exactly two electrons per bond. By interface bond ratio, we mean the (Ga + Se)/(As + Zn) ratio in the interface. Farrell et al. demonstrated that certain GaAs surface reconstructions are optimal in this regard. In the case of GaAs(001) (2 × 4), they showed that As-lean variants of this reconstruction result in a 50:50 interface bond ratio, but the missing row structure for the same surface will give an interface bond ratio of 25:75. They also showed two other surface structures, GaAs(001) *c*(6 × 4) and GaAs(001) (2 × 6), which allow the growth of ZnSe with a 50:50 interface bond ratio. There is some experimental evidence that lends support to this model. Tamargo et al.[101] studied the growth of ZnSe on GaAs(001) and obtained improved crystal quality when the substrate surface reconstruction was GaAs(001) (2 × 4) rather than GaAs(001) (4 × 2). This was attributed to the tendency for development of the As-lean structures in the former case, resulting in a 50:50 interface bond ratio.

These same principles should apply generally to heteroepitaxial growth systems that involve a change in ionicity, although little work has been done along these lines with other material combinations. However, it is clear that surface reconstructions and the resulting interface structures can have far-reaching influence on the properties of heteroepitaxial materials.

4.3 Nucleation

An important difference between homoepitaxy and heteroepitaxy is that heteroepitaxy requires nucleation of a new phase on the substrate surface. Often, the nucleation step greatly influences the morphology and structural properties of heteroepitaxial layers. The nucleation and growth mode, though different, are closely related. Broadly speaking, there are three modes of heteroepitaxial growth;[102] these are the Frank–van der Merwe[103] (FM) (or two-dimensional [2-D], layer-by-layer growth), Volmer–Weber[104] (VW) (or 3-D or island growth), and Stranski–Krastanov (SK) mechanisms. In the case of FM growth, islands of monolayer height coalesce before a new layer can nucleate on top of them. In VW growth, growth proceeds to many atomic layers at discrete islands before these islands merge. In the SK mechanism, the growth is initiated in a layer-by-layer fashion, but islanding commences after the growth of a certain thickness. In all but a few situations, layer-by-layer growth is desirable because of the need for multilayered structures with flat interfaces and smooth surfaces. This requires that the nucleation occur as a single event, on the substrate surface, but not in the gas phase. This section describes the physics and chemistry of nucleation. This treatment begins with homogeneous nucleation, which serves as a useful starting point for the more relevant situation of heterogeneous nucleation. Then heterogeneous nucleation is considered, including the development of a rate equation. Growth modes are treated separately in Section 4.4.

4.3.1 Homogeneous Nucleation

Homogeneous nucleation corresponds to direct condensation out of the gas phase, in the absence of a substrate surface. Though it is rarely observed in practice,* homogeneous nucleation serves as a useful starting point for the discussion of the more important case of heterogeneous nucleation on a substrate.

Homogeneous nucleation[102] may proceed if a condition of supersaturation exists, meaning that the partial pressure P_0 of the nucleating species exceeds the equilibrium vapor pressure P_∞ over the solid phase. In this case, the free energy difference per atom between the vapor and the solid is

$$\Delta g_v = kT \int_{P_0}^{P_\infty} \frac{dP}{P} = -kT \ln\left(\frac{P_0}{P_\infty}\right) = -kT \ln S\,, \tag{4.1}$$

where $S = P_0/P_\infty$ is the supersaturation.

* A notable exception is the epitaxy of silicon from silane, for which special precautions must be taken to avoid gas phase nucleation.

The change in free energy per unit volume of solid is then

$$\Delta g_v = -nkT \ln S, \tag{4.2}$$

where n is the number of atoms per unit volume of the semiconductor.

Under a condition of supersaturation ($S > 1$), embryos will form from the vapor phase. Some embryos will grow by the inclusion of additional material from the vapor phase, if they are large enough so that their growth reduces the overall free energy of the system. Other embryos will be less than the critical embryo size and will shrink by reevaporation. The critical embryo size, above which growth will proceed, may be found by balancing the volume and surface free energy contributions of the embryo. If the surface free energy per unit area of the solid is γ, then the total free energy change for a spherical embryo is given by

$$\Delta G = \frac{4\pi r^3}{3} \Delta g_v + 4\pi r^2 \gamma . \tag{4.3}$$

Figure 4.6 shows the behavior of ΔG qualitatively, for two different values of ΔG_v. With increasing embryo size, ΔG first increases, due to the dominant surface energy term, but then reaches a maximum and finally decreases as the reduction in the free energy due to the phase change prevails. The critical embryo size r_{crit} corresponds to the maximum change in free energy; a larger value of supersaturation (and therefore a more negative ΔG_v) will result in a smaller critical embryo size, as shown in Figure 4.6. The lower curve represents a situation with a higher supersaturation, and thus a more negative value of ΔG_v, resulting in a smaller size for the critical nucleus. In other words $\Delta G_{v1} < \Delta G_{v2}$, so that $r_{\text{crit1}} < r_{\text{crit2}}$ and $\Delta G_{\text{homo1}} < \Delta G_{\text{homo2}}$.

The critical embryo size can be evaluated by setting

$$\left.\frac{\partial \Delta G}{\partial r}\right|_{r=r_{\text{crit}}} = 0, \tag{4.4}$$

which yields

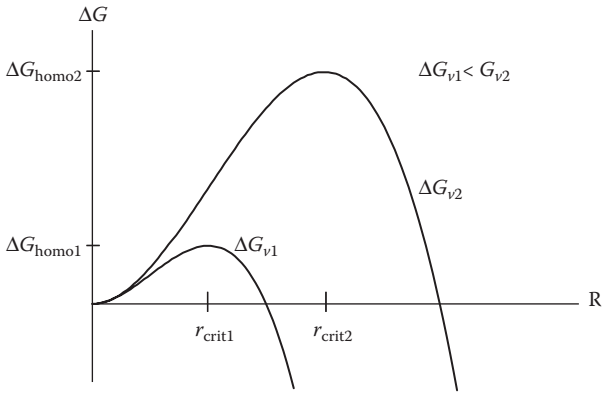


FIGURE 4.6
Total free energy change ΔG for an embryo of the solid phase, as a function of the embryo radius R, for two values of ΔG_v.

$$r_{\text{crit}} = -\frac{2\gamma}{\Delta G_v}. \tag{4.5}$$

Embryos larger than this size will lower their free energy by continuing to grow, whereas those smaller than this size will reduce their free energy by shrinking. The change in free energy evaluated at the critical embryo size is the activation energy for the formation of embryos, which is

$$\Delta G_{\text{homo}} = \frac{16\pi\gamma^3}{3\Delta G_v{}^2}. \tag{4.6}$$

If the nucleation process occurs in an ideal gas, then from the kinetic theory of gases the rate at which atoms (or molecules) will arrive at the surface of a spherical nucleus having the critical radius will be

$$j = \frac{4\pi r_{\text{crit}}^2 P_0}{\sqrt{2\pi mkT}}, \tag{4.7}$$

where:

- m is the atomic (molecular) mass
- k is the Boltzmann constant
- T is the absolute temperature.

Then, assuming Boltzmann statistics, the homogeneous nucleation rate (nuclei formed per unit time in unit volume) can be estimated as

$$R_{\text{homo}} \approx \frac{4\pi r_{\text{crit}}^2 P_0 n}{\sqrt{2\pi mkT}} \exp\left(-\frac{\Delta G_{\text{homo}}}{kT}\right), \tag{4.8}$$

where n is the atomic (molecular) density of the gas. Sometimes the expression above is multiplied by a unitless Zeldovich factor Z, which describes the probability that a critical nucleus will grow rather than shrink.

In the simple model outlined here, the Zeldovich factor was assumed to be unity and the nuclei were assumed to be spherical in shape. Real crystal nuclei will take on more interesting shapes, so that Equation 4.8 will apply only roughly. However, this result shows that the nucleation rate will vary strongly with the supersaturation (and therefore reactant partial pressure) and with the temperature, two important growth parameters.

4.3.2 Heterogeneous Nucleation

Heterogeneous nucleation takes place in the presence of a surface and is more relevant to the case of heteroepitaxy. Here, the presence of the surface changes the situation significantly and can greatly alter the nucleation rate. Heterogeneous nucleation may be considered from a macroscopic point of view,[102] in a fashion paralleling the treatment of the homogeneous case. Atomistic models[105–107] have also been developed, which are applicable to nuclei containing as few as two atoms. In the following subsections, the macroscopic model will be outlined, followed by an atomistic model. Then the case of vicinal substrates

will be described. Through these discussions, *nucleation* will be used to refer to the growth of nuclei on a foreign substrate (as in the initiation of heteroepitaxy) *but also* the nucleation of new clusters of epitaxial material on top of an established layer of this same crystal (second-layer nucleation).

4.3.2.1 Macroscopic Model for Heterogeneous Nucleation

Extending the macroscopic model to the case of heterogeneous nucleation, we find that the presence of a surface tends to increase the nucleation rate. This is because the nuclei may "wet" the substrate, greatly changing its geometry. Suppose γ_e and γ_s represent the surface free energies of the epitaxial crystal and substrate, respectively, and γ_i is the epitaxial layer–substrate interfacial free energy. The epitaxial material will not wet the substrate if $\gamma_i > (\gamma_e + \gamma_s)$, because this would be accompanied by an overall increase in the free energy of the system. On the other hand, complete wetting is expected if $(\gamma_e + \gamma_i) < \gamma_s$ (the epitaxial deposit will spread out to maximize the area of the interface). For all other situations, consideration of force balance on the boundary of the embryo leads to the expectation of partial wetting with a contact angle θ, where

$$\gamma_s = \gamma_i + \gamma_e \cos\theta. \tag{4.9}$$

or

$$\theta = \cos^{-1}\left(\frac{\gamma_s - \gamma_i}{\gamma_e}\right). \tag{4.10}$$

These three types of situations are shown schematically in Figure 4.7.

Assuming the embryos to be sphere segments with the appropriate contact angle θ, we can follow a development paralleling that for the homogeneous case. The free energy change upon formation of the embryo, associated with the surface and interfacial energies, will be $A_e\gamma_e + A_i(\gamma_i - \gamma_s)$, where A_e and A_i are the areas of the embryo surface (with the gas phase) and interface (with the substrate), respectively. Therefore, the total free energy change associated with creation of the embryo on the substrate surface will be

$$\Delta G = \frac{\pi r^3}{3}(1-\cos\theta)^2(2+\cos\theta)\Delta G_v + \pi r^2(\gamma_i - \gamma_s)\sin^2\theta + 2\pi r^2(1-\cos\theta)\gamma_e, \tag{4.11}$$

where r is the radius of curvature for the (truncated sphere) embryo.

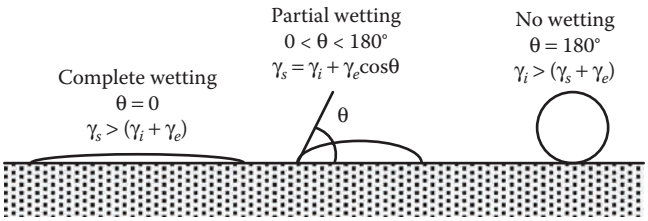


FIGURE 4.7
Wetting of a flat substrate by an epitaxial deposit. (Adapted from S. K. Ghandhi, *VLSI Fabrication Principles, Silicon and Gallium Arsenide*, 2nd ed. [New York: Wiley, 1994]. Used with permission.)

The free energy change reaches a maximum value of

$$\Delta G_{\text{het}} = \Delta G_{\text{homo}} \left[\frac{(2+\cos\theta)(1-\cos\theta)^2}{4} \right]$$

$$= \frac{16\pi\gamma_e^{\,3}(2+\cos\theta)(1-\cos\theta)^2}{12\Delta G_v^{\,2}} \quad . \tag{4.12}$$

The rate for heterogeneous nucleation will then be

$$R_{\text{het}} \approx C_1 \exp\left(-\frac{\Delta G_{\text{het}}}{kT} \right). \tag{4.13}$$

Here, the prefactor C_1 will be different from the case of homogeneous nucleation, due to the reduction of the embryo surface area by wetting. Like before, the nucleation rate will depend very strongly on the supersaturation and the temperature. However, the change in the critical free energy change can drastically increase the nucleation rate for given conditions of supersaturation and temperature, compared with the homogeneous case. This is fortunate, for it allows heteroepitaxial growth to occur under conditions that suppress gas phase nucleation. However, layer-by-layer growth requires that the heterogeneous nucleation proceed at a slow rate of one event per monolayer.

4.3.2.2 Atomistic Model

The macroscopic model for heterogeneous nucleation is based on the macroscopic properties, such as the surface and interfacial free energies. In some cases, heterogeneous nucleation may occur with nuclei containing as few as two atoms. In such cases, an atomistic model[105–107] for nucleation is more relevant.

In developing an atomistic model for heterogeneous nucleation, it is assumed that atoms arrive at a flat surface with an impingement rate of F (atoms per unit area per unit time*). This incident flux gives rise to a concentration of adatoms (per unit area) on the surface equal to n_1. There will also be unstable clusters of two or more atoms on the surface. Here, n_j will be used to denote the concentration (per unit area) of clusters containing j atoms. Unstable clusters can reduce the free energy of the system by shrinking. However, there will also be a concentration of stable clusters on the surface. These are large clusters that reduce the total free energy of the system by growing. If the critical cluster size contains i atoms, then all clusters containing more than i atoms will be stable. Here, n_x denotes the concentration of stable clusters on the surface, and w_x is the average number of atoms in a stable cluster ($w_x > i$).

Suppose atoms arrive at the flat surface with a rate F (atoms per unit area per unit time). The interaction of the adatoms and surface clusters with the gas phase can be illustrated schematically as in Figure 4.8. Adatoms may reevaporate (with a time constant τ_a), combine with other adatoms or unstable clusters, be captured by a critical cluster (nucleation, with a time constant τ_n), or be captured by a stable cluster (with a time constant τ_c). Any critical cluster of i atoms that succeeds in capturing one more adatom will become a stable cluster.

* In the literature, the incident flux and cluster concentrations are sometimes given in atoms per second and absolute numbers, respectively. Here, these quantities will be given on a per unit area basis, so that R has units of atoms per unit area per unit time, and n_1 has units of adatoms per unit area.

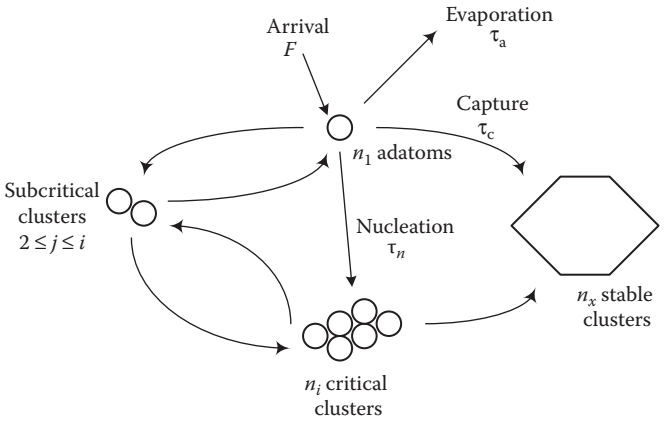


FIGURE 4.8
Schematic diagram of the interactions between the adatoms and surface clusters with the gas phase. All arriving atoms condense with a rate F. This gives rise to an adatom population n_1. Adatoms may reevaporate (with a time constant τ_a), combine with another adatom or unstable cluster (nucleation, with a time constant τ_n), or be captured by a stable cluster (with a time constant τ_c). Any critical cluster of i atoms that succeeds in capturing one more adatom will become a stable cluster. (Adapted from J. A. Venables, *Phys. Rev. B*, 36, 4153 [1987]. With permission. Copyright 1987, American Physical Society.)

In steady state, adatoms are emitted and accepted at equal rates by unstable clusters, with no net effect on the populations of the adatoms or the unstable clusters.

The rate equations for this system have been derived by Stowell and Hutchinson[108] and Stowell.[109] They are the following:

$$\frac{dn_1}{dt} = F - \frac{n_1}{\tau_a} - \frac{d\left(n_x w_x\right)}{dt}, \tag{4.14}$$

$$\frac{dn_j}{dt} = 0, \left(2 \le j \le i\right), \tag{4.15}$$

and

$$\frac{dn_x}{dt} = \sigma_i D n_1 n_i - 2n_x \frac{dZ}{dt}. \tag{4.16}$$

Equation 4.14 gives the time rate of change of the adatom concentration, in which the three terms represent condensation, reevaporation, and diffusive capture by stable clusters.* Equation 4.15 stems from the assumption that the populations of unstable clusters are constant with time; this is true if the growth occurs near equilibrium. Equation 4.16 gives the time rate of change of the concentration of stable clusters. Here, the first term represents the creation of new stable clusters (the *nucleation rate*, in nuclei per unit area per unit time) by the diffusive capture of adatoms (with concentration n_1) by critical clusters (with population n_i). D is the diffusivity of adatoms and σ_i is the (unitless) capture number for the critical size clusters. The second term in Equation 4.16 represents the coalescence of stable clusters, and Z is the fraction of the surface covered by stable clusters ($0 \le Z \le 1$).

* The rate of nucleation is numerically unimportant here, for the purpose of determining n_1, and was neglected.

Three other basic relationships are needed to solve Equations 4.14 through 4.16 and determine the nucleation rate. First, Equations 4.14 and 4.16 are coupled through

$$\frac{d\left(n_x w_x\right)}{dt} = \frac{n_1}{\tau_n} + \frac{n_1}{\tau_c} + FZ. \tag{4.17}$$

Moreover, in steady state, dn_1/dt, so that

$$n_1 = F\tau\left(1 - Z\right), \tag{4.18}$$

where $\tau^{-1} = \tau_a^{-1} + \tau_n^{-1} + \tau_c^{-1}$ and $\tau_c^{-1} = \sigma_x D n_x$, where σ_x is the effective capture number (unitless) for stable clusters. Second, the substrate coverage Z is related to the number of atoms in the stable clusters by

$$\frac{dZ}{dt} = N_a^{-1}\frac{d\left(n_x w_x\right)}{dt}, \tag{4.19}$$

where N_a is the areal density of atoms in the stable clusters. Assuming monolayer clusters, $N_a^{-1} = \Omega^{2/3}$, where Ω is the atomic volume. Third, the relationship between the populations of critical clusters and adatoms, in the case of a relatively high supersaturation, is given by

$$\frac{n_i}{N_0} = \left(\frac{n_1}{N_0}\right)^i C_i \exp\left(\frac{E_i}{kT}\right), \tag{4.20}$$

where:

N_0 is the atomic density in the substrate crystal (atoms per unit area)
E_i is the free energy change associated with the critical size cluster
C_i is a constant

This equation can be generalized to account for more than one configuration of critical clusters, if there is a low supersaturation.

Upon solution of the Equations 4.14 through 4.20, we obtain the normalized density of stable clusters (nuclei), assuming monolayer islands and negligible evaporation, as

$$\frac{n_x}{N_0} = C\eta\left(\frac{F}{N_0\nu}\right)^{\frac{i}{i+2}} \exp\left[\frac{E_i + iE_d}{\left(i+2\right)kT}\right], \tag{4.21}$$

where:

C and η are constants
ν is the effective surface vibration frequency (~10^{11} to 10^{13} s^{-1})

The nucleation rate, also assuming monolayer islands and negligible evaporation, is

$$J = \sigma_i D n_1 n_i. \tag{4.22}$$

The atomistic model outlined here has been applied to numerical calculations of the stable cluster densities in several material systems.[107] These calculations provide reasonable agreement to experimental results for several combinations of metal-on-metal,[110]

metal-on-Si,[111] and rare gases on amorphous carbon.[112] This and similar atomistic models for nucleation have also been applied to the analysis of second-layer nucleation in semiconductor heteroepitaxy.[113] This has made it possible to predict surface roughening associated with a transition from a layer-by-layer growth mode to an SK (layer plus islands) growth mode.

An atomistic approach to nucleation is appropriate if the critical nuclei comprise only a few atoms, for in this case the macroscopic surface and interfacial energies are inapplicable. In the limit of large critical cluster size, however, the atomistic and macroscopic models should converge. The proof of this is not trivial, however, and so at the present time no unified model has emerged.

The atomistic model described above is convenient in that it is analytical in nature. But it is also completely deterministic, and cannot account for the statistical nature of surface atomic processes such as diffusion. Completely stochastic atomistic models have also been implemented in the form of molecular dynamic (MD)[114,115] or kinetic Monte Carlo (KMC)[116,117] numerical simulations, which address this issue, but which are beyond the scope of this book.

4.3.2.3 Vicinal Substrates and Step-Flow Growth

In the case of heteroepitaxy of semiconductors, vicinal (tilted) substrates are often utilized. An example is GaAs(001) 2°→ [110]. The surface of a vicinal substrate comprises low-index terraces separated by monolayer or bilayer steps, which have a separation determined by the offcut angle. Generally, there will also be kinks (jogs in the steps) whose density will depend on the direction of the offcut. Here, heterogeneous nucleation of the dissimilar material of the vicinal substrate will occur preferentially at the kinks or steps, due to the modified value of ΔG_{het} compared with the open terraces. Therefore, growth on a vicinal substrate by the advancement of steps (*step-flow* growth) may be possible at a supersaturation too low to result in nucleation on a flat substrate, and may allow improved crystal quality and the suppression of islanding.

4.4 Growth Modes

There are three broad classifications for the growth modes for heteroepitaxy.[102] These are the Frank–van der Merwe[103] (FM) (or 2-D growth), Volmer–Weber[104] (VW) (or 3-D growth), and Stranski–Krastanov (SK)* mechanisms. FM growth on a flat substrate† is characterized by the nucleation of a new monolayer and its growth to cover the substrate, followed by the

* This growth mode was named Stranski–Krastanov by Bauer and Poppa.[118] This came about as a result of a calculation made by Stranski and Krastanov.[119] They showed that, for a monovalent ionic crystal condensing onto a divalent substrate, the second layer of condensate has weaker bonding than the substrate surface layer, even though the first layer of condensate has stronger bonding. It is possible that this phenomenon could result in the mode we have come to know as SK.

† In the case of a vicinal substrate, the surface comprises a number of flat terraces separated by monolayer steps. Here, the nucleation of new layers is unnecessary. Instead, growth proceeds by the advancement of steps and this is called "step-flow growth." In either case (layer-by-layer or step-flow growth), the epitaxial layer retains the surface smoothness of the starting substrate.

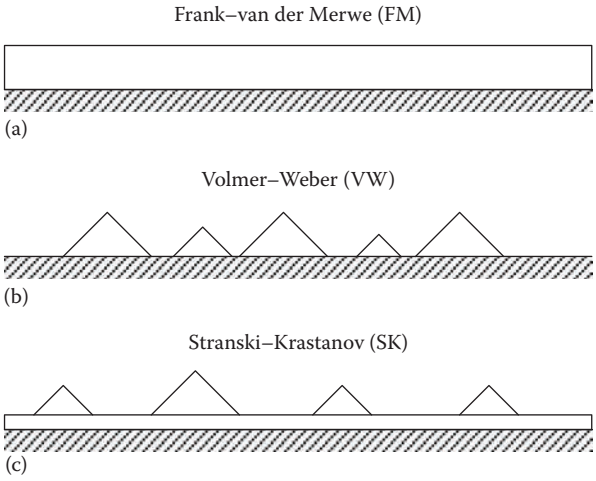


FIGURE 4.9
(a) FM, (b) VW, and (c) SK growth modes for heteroepitaxy.

nucleation of the next layer. This growth mode is therefore referred to as "layer-by-layer" growth. VW growth involves the development of isolated islands on the substrate, followed by their growth and coalescence. This coalescence process results in a rough surface, with a root mean square (rms) roughness comparable to the mean distance between islands. In the SK mechanism, the initial growth proceeds in a layer-by-layer fashion but becomes 3-D in nature after the growth of a certain critical layer thickness. (It should be emphasized that this is not the same as the critical layer thickness for lattice relaxation, although the two may be comparable in size and are sometimes used interchangeably.) The FM, VW, and SK growth modes are illustrated schematically in Figure 4.9.

The important distinction between 2-D growth and the other modes is that in a 2-D growth mode either (1) a monolayer completes before a new one nucleates or (2) if multiple nucleation events occur on a flat substrate surface, then the monolayer islands coalesce before additional growth occurs on top of them. Therefore, VW and SK growth, which involve islanding, are sometimes collectively called "multilayer growth," referring to the fact that islands grow beyond one monolayer thickness before coalescing.

In all but a few situations, 2-D growth (either layer-by-layer or step-flow growth) is desirable because of the need for multilayered structures with flat interfaces and smooth surfaces. A notable exception is the fabrication of quantum dot devices, which require VW or SK growth of the dots. Even here, though, it is desirable for the other layers of the device to grow in a 2-D mode. For example, Guha et al. found that, in the case of ZnSe/GaAs(001) heteroepitaxy, layers that grew by a 2-D growth mode exhibited an order of magnitude lower threading dislocation density than layers that grew by the coalescence of islands.[120] In all cases of heteroepitaxy, however, it is important to be able to control the nucleation and growth mode.

The growth modes in heteroepitaxy have been considered extensively based on thermodynamic models.[121–123] Along these lines, Daruka and Barabási[122,123] developed an equilibrium phase diagram that identifies the growth mode as a function of the lattice mismatch strain and the average thickness of the deposit. At the same time, it has also been established that kinetic factors often play an important role in establishing the growth mode, if the growth proceeds with a large supersaturation.[124] Of these, the most important are

the surface diffusivity, the energy barrier to diffusion at steps, and the growth rate. Based on kinetic considerations, Tersoff et al.[113] showed that there is a critical island size for the achievement of layer-by-layer growth. While these aspects of heteroepitaxy are far from completely understood, it is becoming clear that kinetic factors provide an opportunity for controlling the growth mode. An especially interesting aspect of this involves the use of surfactants. While the behavior of surfactants in heteroepitaxy is not yet entirely clear, it is known that surfactants can in some cases modify the growth mode. In the following sections, equilibrium considerations will first be presented, including the development of a general growth mode phase diagram. The role of misfit dislocations will be considered, with application to SiGe/Si. This will be followed by a brief consideration of a kinetic model, and a development of the conditions necessary for layer-by-layer growth. The possible roles of surfactants are considered in Section 4.6.

4.4.1 Growth Modes in Equilibrium

In the classical theory, the mechanism of heterogeneous nucleation is dictated by the surface and interfacial free energies for the substrate and epitaxial crystal.[102] The energy criteria are stated in terms of $\Delta\gamma$, the areal change in free energy associated with covering the substrate with the epitaxial layer, not including the bulk free energy of the epitaxial crystal. Then, if γ_e and γ_s are the surface free energies of the epitaxial layer and substrate, respectively, and γ_i is the interfacial free energy for the epitaxial–substrate interface, then

$$\Delta\gamma = \gamma_e + \gamma_i - \gamma_s. \tag{4.23}$$

If minimum energy dictates the mode for nucleation and growth, the prevalent mechanism will be 2-D for $\Delta\gamma < 0$ and 3-D for $\Delta\gamma > 0$. Often, the interfacial contribution can be neglected in comparison to the surface energy terms. If this is the case, then 2-D growth is expected for $\gamma_e < \gamma_s$ (the epitaxial layer will wet the substrate), but 3-D growth will occur if $\gamma_e > \gamma_s$. However, even in the case of a wetting epitaxial layer ($\gamma_e < \gamma_s$), the existence of mismatch strain can cause islanding after the growth of a few monolayers. This is because the strain energy in the coherent epitaxial layer increases in direct proportion to the thickness. At some point, it becomes energetically favorable to create islands that can relieve some of the mismatch strain by relaxation at the sidewalls. Therefore, SK growth can be expected in the case of a wetting epitaxial layer unless the lattice mismatch strain is quite small. Whereas it is clear that the VW growth mode is to be expected for a nonwetting epitaxial layer, the behavior of a wetting deposit is more complex and warrants further consideration.

In order to elucidate this behavior, Daruka and Barabási[122,123] investigated the growth of a lattice-mismatched, wetting epitaxial layer on a foreign substrate, and created an equilibrium phase diagram that can help predict the growth mode for heteroepitaxy. In the development of their model, they assumed the growth of a wetting epitaxial layer B on a substrate A, with a thickness of H monolayers and a lattice mismatch f. The total deposit is distributed between the wetting layer with a thickness of n_1 monolayers, stable islands with an average thickness of n_2 monolayers, and large, ripened islands having an average thickness of $H = n_1 - n_2$ monolayers. Both stable and ripened islands were assumed to be square pyramids with a fixed aspect ratio; this aspect ratio corresponds to crystal faces for which the facet energy has a minimum as determined on Wulff's plot.[125]

In their calculations, Daruka and Barabási neglected evaporation and considered the free energy per interfacial lattice site (effectively per unit area), $f = u - T_s$, where u is the

internal energy density, T is the temperature, and s is the entropy density. Furthermore, they assumed that the entropy contribution is negligible, which has been shown to be appropriate for lower temperatures,[123] so that $f \approx u$. The average free energy per lattice site for the combination of wetting layer and islands was calculated to be

$$u = n_1 E_{ml} + n_2 E_{isl} + (H - n_1 - n_2) E_{rip}, \tag{4.24}$$

where:

E_{ml} is the energy per monolayer in the strained wetting layer
E_{isl} is the free energy per monolayer in the pyramidal islands
E_{rip} is the energy per monolayer in the ripened islands

The energy density of the coherently strained wetting layer may be calculated to first order as $G = Cf^2 - \Phi_{AA}$, where C is a constant that depends on the biaxial modulus and the monolayer thickness, and $-\Phi_{AA}$ is the energy of an AA atomic bond. Daruka and Barabási accounted for the energy of the interfacial AB bonds, $-\Phi_{AB}$, and also the fact that the binding energies of A atoms will be modified close to the interface.[126,127] Including these contributions, the energy density in the wetting layer was assumed to be

$$E_{ml} = \int_0^{n_1} \left\{ G + \Delta \left[U(1-n) + U(n-1) e^{-(n-1)/a} \right] \right\} dn, \tag{4.25}$$

where the parameter a specifies the effective range for the interatomic forces (Daruka and Barabási assumed that $a = 1$ in their calculations), $\Delta = \Phi_{AA} - \Phi_{AB}$, and $U(x)$ is the unit step function:

$$U(x) = \begin{cases} 0; & x < 0, \\ 1; & x > 0. \end{cases} \tag{4.26}$$

Daruka and Barabási noted that the exact form of E_{m1} would not change the qualitative features of the overall behavior.

The free energy per monolayer in the pyramidal islands was calculated using

$$E_{isl} = gCf^2 - \Phi_{AA} + E_0 \left[-\frac{2}{x^2} \ln\left(x\sqrt{e} \right) + \frac{\alpha}{x} + \frac{\beta}{x^{3/2}} \right], \tag{4.27}$$

where g is a form factor that expresses the reduction in the strain energy of the islands compared with the continuous wetting layer and $0 < g < 1$. The normalized island size is $x = L/L_0$, where L is the length of a side of the pyramidal island and L_0 is a characteristic length as defined by Shchukin and coworkers.[128] The three terms in the square brackets arise because of the faceting of the islands. With nonzero facet surface energy, compressive forces will develop at the facet edges, resulting in a component of stress energy in the islands,[129] which Daruka and Barabási have called the "homoepitaxial stress." The first term in the square brackets corresponds to this homoepitaxial stress contribution. The second term in the square brackets is associated with the cross-term from the interaction of the lattice mismatch stress and the homoepitaxial stress, and also the facet energy, with has the same x dependence. Therefore, $\alpha = p(\gamma - f)$, where p and γ are material constants. The third term in the square brackets represents the energy from the island–island stress interaction.

The free energy per monolayer in the ripened islands is

$$E_{\text{rip}} = gCf^2 - \Phi_{AA}. \tag{4.28}$$

Daruka and Barabási used the model outlined here to calculate phase diagrams using various values of the material parameters. To do this, they minimized the energy with respect to n_1, n_2, and x. The detailed procedure has been described by Daruka.[123]

The Daruka and Barabási equilibrium model involves a number of material parameters (a, C, E_0, Φ_{AA}, Φ_{AB}, g, p, γ, and b), some of which are either semiempirical or not known with a high degree of accuracy. Therefore, the question arises as to whether the choice of material parameters will change the very nature of the phase diagram. However, Daruka[123] showed that, for any set of material parameters, the resulting phase diagram would assume one of four basic topologic forms. Moreover, Daruka and Barabási developed a general phase diagram that incorporates the features of all four classes, as shown in Figure 4.10, using the parameters $a = 1$, $C = 40E_0$, $\Phi_{AA} = E_0$, $\Phi_{AB} = 1.27E_0$, $g = 0.7$, $p = 4.9$, $\gamma = 0.3$, and $b = 10$. Here seven distinct phase regions are seen, corresponding to FM, SK, VW, or ripening (R) behavior. There are two SK phase regions, SK_1 and SK_2. In both cases, islands coexist with a wetting layer; however, as will be explained in the following, the behavior with increasing growth time is different in the two cases. There are also three ripening phases:

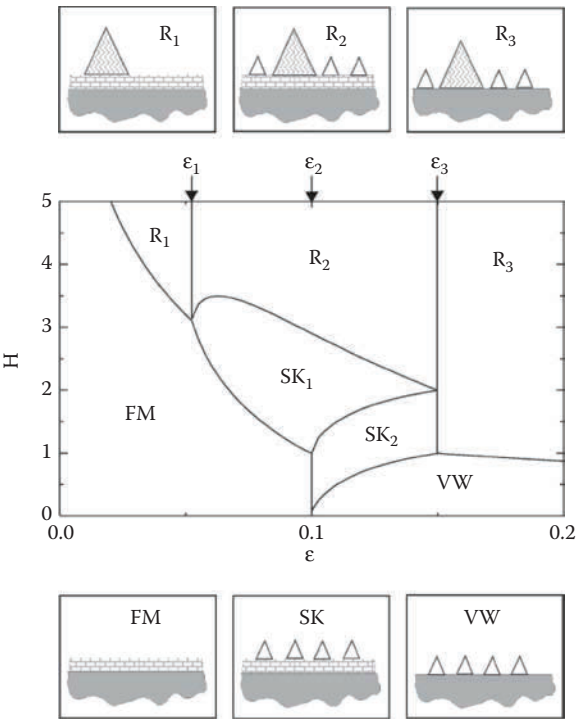


FIGURE 4.10
Equilibrium phase diagram for heteroepitaxy of a wetting material A on a substrate B. H is the average thickness in monolayers, and ε is the lattice mismatch strain. The small panels show the morphology of the growing film for each of the phase regions. The small open triangles represent stable islands, whereas the large shaded triangles denote ripened islands. The parameters used to calculate the phase diagram were $a = 1$, $C = 40E_0$, $\Phi_{AA} = E_0$, $\Phi_{AB} = 1.27E_0$, $g = 0.7$, $p = 4.9$, $\gamma = 0.3$, and $b = 10$. (Reprinted from I. Daruka and A.-L. Barabási, *Phys. Rev. Lett.*, 79, 3708 [1997]. With permission. Copyright 1997, American Physical Society.)

R_1, R_2, and R_3. The R_1 phase exhibits ripened islands along with a wetting layer. The R_2 phase exhibits stable islands as well. The R_3 phase is characterized by the presence of both ripened and stable islands, but no wetting layer.

Growth phases involving ripened islands are generally undesirable. Based on thermodynamic considerations, the ripened islands are predicted to have infinite size and vanishing density on the surface. In a real epitaxial growth process, ripening islands will have finite size and density, as determined by the kinetics of their growth. The important distinction is that they will be unstable and there will exist a driving force for their growth with time by the process of Ostwald ripening.[130–132] Therefore, stable islands are needed for device fabrication if additional high-temperature processing will be used following their growth.

In a typical heteroepitaxial growth process, the lattice mismatch strain f (ε in the notation of Daruka and Barabási) is fixed, but the extent of the deposit H (in monolayers) increases with time. Based on the phase diagram contained in Figure 4.10, we can understand four cases of such a process, which will be outlined in the following.

4.4.1.1 Regime I ($f < \varepsilon_1$)

Suppose the lattice mismatch strain is small ($f < \varepsilon_1$, with ε_1 indicated in Figure 4.10); one possible value of lattice mismatch is indicated by the arrow at the bottom of the phase diagram. In this case, the initial growth will occur with an FM mode. After the deposition of a certain thickness, however, we expect a transition to the R_1 phase, and so ripened islands will grow on the initial wetting layer. In this phase region, the wetting layer thickness does not increase but stays constant, so that the newly deposited material contributes only to the formation of ripened islands. The energy is minimized when the ripened islands approach infinite size with zero density.*

4.4.1.2 Regime II ($\varepsilon_1 < f < \varepsilon_2$)

If the lattice mismatch strain is increased somewhat ($\varepsilon_1 < f < \varepsilon_2$, with ε_1 and ε_2 indicated in Figure 4.10), the initial growth still exhibits a 2-D nature. As the average thickness of the deposit is increased, however, we can expect a transition to an SK mode (phase SK_1) in which stable islands with finite size and density grow on top of the initial wetting layer. A further increase in the growth time will give rise to the appearance of ripened islands along with the stable islands (phase region R_2). As in regime I, equilibrium considerations predict the ripened islands will have infinite size and zero density.

4.4.1.3 Regime III ($\varepsilon_2 < f < \varepsilon_3$)

Consider a large value of the lattice mismatch strain ($\varepsilon_2 < f < \varepsilon_3$, with ε_2 and ε_3 indicated in Figure 4.10). Here, the lattice mismatch is too large to permit 2-D growth. Instead, the initial growth occurs in a VW mode, with separated, stable islands and the absence of a wetting layer. However, with increasing growth time, the SK_2 phase is encountered. As expected, the SK_2 phase is characterized by stable islands and a wetting layer. But in this case, the islands grow first, followed by the wetting layer, which fills in the area separating them. As growth proceeds, the wetting layer increases its thickness, but the stable islands remain fixed in size. This continues until the SK_1 phase boundary is encountered. Then, in the SK_1 phase region, continued growth increases the number of stable islands while

* Of course, kinetic considerations would predict islands with a finite size and finite density.

the wetting layer remains at constant thickness. Eventually, the growth proceeds in the R_2 phase, in which ripened islands grow from the additional material.

4.4.1.4 Regime IV ($f > \varepsilon_3$)

For very high values of the lattice mismatch strain ($f > \varepsilon_3$, with ε_3 indicated in Figure 4.10), the initial growth occurs with a VW mode, followed by the growth of ripened islands (phase region R_3). A wetting layer never forms, and so a continuous heteroepitaxial layer will not be achieved in this case.

In all four regimes of mismatch described above, equilibrium theory predicts the growth of infinitely large ripened islands, with vanishing density. This cannot occur in a real growth situation, in which the growth and surface diffusion processes occur at finite rates. So, although the equilibrium considerations outlined in this section provide guidance in terms of the driving forces and the direction in which growth will proceed, the kinetic considerations will dictate the density and size of ripening islands and perhaps also the "stable islands." An important result of this is that the growth morphology can be influenced by changing the growth temperature or growth rate, or by the introduction of surfactant species, which can significantly modify the surface diffusion.

4.4.1.5 Equilibrium Growth Modes in SiGe/Si

In addition to the considerations outlined above in the theory of Daruka and Barabási, it has been shown by Nakajima et al.[133] that the presence of misfit dislocations influences the equilibrium growth mode in the case of SiGe/Si. In their work, Nakajima et al. considered the strain, surface, and interface energies in SiGe/Si(001) and SiGe/Si(111) heterostructures both with and without misfit dislocations for the FM, SK, and VW growth modes. Through this analysis, they were able to determine the free energy for each growth mode as a function of Ge mole fraction and layer thickness, and therefore determine the equilibrium growth mode phase diagrams for both of these crystal orientations, as illustrated in Figures 4.11 and 4.12 for the (111) and (100) orientations, respectively. The results show that layer-by-layer FM growth is favored for lower Ge concentrations up to approximately 50%. With larger Ge concentrations, SK growth or island formation on a wetting layer was favored. Although the VW growth mode is energetically stable in high-concentration layers with several atomic layers, Nakajima et al. concluded that this mode might not be observed in actual growth because the VW region lies above the SK region of the phase diagram. An important difference between the two orientations is that the regions of SK and VW growth are larger for the (111) case because the strain energy of the (111) face is greater than that of the (100) face. The SK and VW regions are narrowed in the presence of misfit dislocations because of the release of strain energy.

4.4.2 Growth Modes and Kinetic Considerations

Equilibrium considerations dictate that mismatched heteroepitaxial material will usually grow in a VW or SK mode, with a rough surface, unless the epitaxial layer wets the substrate and the lattice mismatch is small. On the other hand, heteroepitaxial growth may occur far from equilibrium (i.e., with a large supersaturation). In such a case, kinetic factors provide an opportunity to tailor the growth mode and morphology. The most important such factors influencing the growth mode and morphology are the surface diffusivity and growth rate (or source flux). In addition to these controllable factors, the diffusion barrier

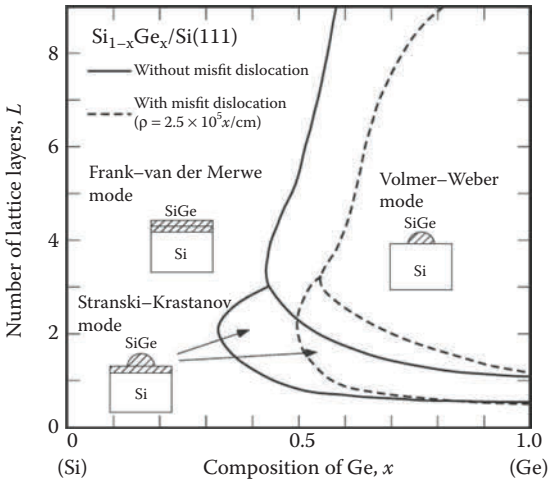


FIGURE 4.11
Thickness composition phase diagram of the FM, SK, and VW growth modes for SiGe/Si (111) without (solid curves) and with (dashed curves) misfit dislocations. (Reprinted from K. Nakajima et al., *J. Cryst. Growth*, 260, 372 [2004]. With permission. Copyright 2004, Elsevier.)

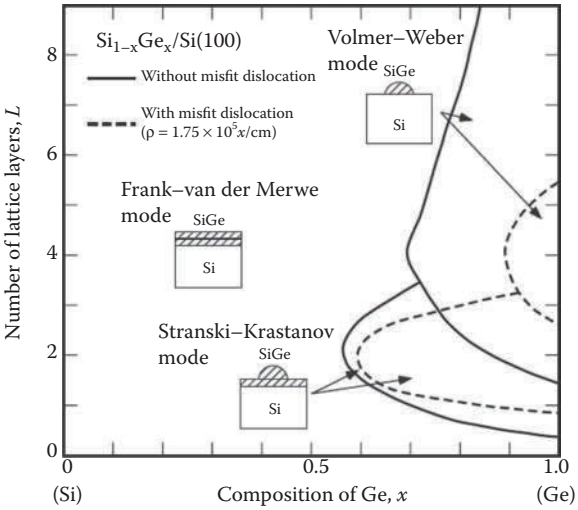


FIGURE 4.12
Thickness composition phase diagram of the FM, SK, and VW growth modes for SiGe/Si (100) without (solid curves) and with (dashed curves) misfit dislocations. (Reprinted from K. Nakajima et al., *J. Cryst. Growth*, 260, 372 [2004]. With permission. Copyright 2004, Elsevier.)

at step edges (the Ehrlich–Schwoebel barrier[134,135]) may also play a role in determining the growth mode. Consideration of kinetics shows that it is possible to tailor the growth conditions (through the temperature, growth rate, or introduction of surfactants) in such a way as to obtain layer-by-layer growth. An intriguing discovery is the existence of reentrant epitaxy,[136,137] in which the growth mode observed at high and low temperatures differs from that found at the middle range of temperatures. Just as interesting is the finding by many workers that the inclusion of a surfactant can alter the growth mode, thereby either inhibiting or promoting island growth.

It should also be possible to design the growth process in such a way as to control the size and density of islands in the SK or VW growth modes. In this section, the condition for layer-by-layer growth will be developed. The control of islanding in heteroepitaxy, also known as "self-assembly," is considered in Section 4.7.

In considering the kinetic factors controlling the growth mode, the problem is to find the conditions that give rise to layer-by-layer growth, rather than the growth of isolated 3-D clusters. Or, stated differently, the problem is to find the conditions under which a new layer will nucleate on a monolayer island before coalescence (so-called second-layer nucleation), which will give rise to kinetic roughening.

Tersoff, Denier van der Gon, and Tromp[113] (TDT) derived the critical island size for layer-by-layer growth by a consideration of this second-layer nucleation process.* In their model, TDT assumed the existence of circular, monolayer islands with uniform radius. They found the rate of second-layer nucleation on top of these islands using classical atomistic nucleation theory, by solving the diffusion equation for adatoms with a constant growth flux (MBE conditions). This model is summarized in what follows.

Based on an atomistic approach, TDT assumed the nucleation rate to be

$$\omega \approx DN_0^2 \left(\frac{n_1}{N_0} \right)^i = DN_0^2 \eta^i, \tag{4.29}$$

where:

- i is the number of atoms in the smallest stable island
- D is the diffusivity for surface atoms
- N_0 is the surface atomic density (atoms per unit area)
- n_1 is the surface concentration of adatoms (per unit area)

The normalized (dimensionless) adatom density is $\eta = n_1/N_0$.

Consider the growth of a heteroepitaxial layer with an incident flux of atoms F. This could correspond directly to the case of MBE, but can be applied to VPE processes by a simple extension. The diffusion equation for the resulting adatoms on the surface is

$$\frac{d\eta}{dt} = D\nabla^2\eta + \frac{F}{N_0}. \tag{4.30}$$

To investigate the possibility of second-layer nucleation, we consider the growth on an existing monolayer island having circular geometry. In this case, the steady-state solution to the diffusion equation for adatoms on the top of the island is

$$\eta = \eta_0 - \frac{F}{4DN_0} r^2, \tag{4.31}$$

where r is the distance from the center of the island. The boundary condition at the island edge is $d\eta/dr + \eta\alpha\sqrt{N_0} = 0$, where α represents the probability that an adatom, upon reaching the island edge, will jump down over the edge in unit time, divided by the rate for hops within the area of the terrace. If there is an energy barrier E_s for hopping over the edge of the island, then $\alpha = C_1 \exp\left[-(E_s - E_d)/kT\right]$, where C_1 is a constant, E_s is the

* Here, second-layer nucleation refers generally to formation of stable nuclei on top of an established island.

Ehrlich–Schwoebel barrier,[134,135] and E_d is the activation energy for surface diffusion on top of the island. Based on the boundary condition above, the constant in Equation 4.31 may be evaluated as

$$\eta_0 = \frac{F}{4DN_0}\left(R^2 + RL_\alpha\right), \tag{4.32}$$

where $L_\alpha \equiv 2/\left(\alpha\sqrt{N_0}\right)$ and R is the average island radius. (In this simple model, the islands are assumed to have uniform size equal to the average size.)

The rate of nucleation on top of the island, in nuclei per unit time, can be found by integrating over the island area:

$$\Omega = \int_0^R \omega 2\pi r dr = \frac{\pi D}{i+1}\left(\frac{F}{2D}\right)^i N_0^{2-i}\left[\left(R^2 + RL_\alpha\right)^{i+1} - \left(RL_\alpha\right)^{i+1}\right]. \tag{4.33}$$

TDT considered two limiting cases. In case 1, the Ehrlich–Schwoebel barrier can be neglected, so $\alpha \approx 1$ and $L\alpha \ll R$, giving

$$\Omega_1 = \frac{\pi D}{i+1}\left(\frac{F}{2D}\right)^i N_0^{2-i} R^{2i+2}. \tag{4.34}$$

In case 2, the Ehrlich–Schwoebel barrier is much greater than kT, so that $\alpha \to 0$ and

$$\Omega_2 = \pi D\left(\frac{F}{4D}\right)^i N_0^{2-i} R^{i+2} L_\alpha^i. \tag{4.35}$$

Case 2 is most relevant to the consideration of second-layer nucleation and can be used to determine the fraction of islands f that have nucleated a second layer on their top. The time rate of increase for the fraction of islands experiencing second-layer nucleation is given by

$$\frac{df}{dt} = \Omega\left(1 - f\right). \tag{4.36}$$

TDT assumed that the growth of the islands with time can be described by

$$R^2 = \frac{FL_n^2}{N_0}t, \tag{4.37}$$

where πL_n^2 is the area per island (the reciprocal of the areal density of the nuclei) so that L_n is approximately the separation between islands. Then the fraction of islands with nuclei on top will be

$$f = 1 - \exp\left[-\left(R/R_c\right)^m\right], \tag{4.38}$$

where:

R_c is the critical island size (radius) for the transition from FM to SK growth
m is a unitless parameter that depends on the critical cluster size.

For case 1, with a negligible Ehrlich–Schwoebel barrier, $m = 2i + 4$ and

$$R_{c1} = \left[(i+1)(2i+4) \left(\frac{2L_n^2}{\pi} \right) \left(\frac{4D}{F} \right)^{i-1} N_0^{i-3} \right]^{1/(i+4)}. \tag{4.39}$$

For case 2, with ($L_\alpha \gg R$), $m = i + 4$ and

$$R_{c2} = \left[(i+4) \left(\frac{2L_n^2}{\pi} \right) \left(\frac{4D}{F} \right)^{i-1} L_\alpha^{-i} N_0^{i-3} \right]^{1/(i+4)}. \tag{4.40}$$

Thus, once the critical island size R_c has been established based on Equation 4.39 or 4.40, the growth mode for a wetting layer with an island size R is predicted as follows:

$$R < R_c \rightarrow FM \tag{4.41}$$

and

$$R > R_c \rightarrow SK. \tag{4.42}$$

Essentially, if ($R > R_c$), there will be new nucleation on the islands before they coalesce. This will give rise to undesirable surface roughening in the case of homoepitaxy or heteroepitaxy. This is illustrated in Figure 4.13, which shows the fraction of islands experiencing second-layer nucleation versus the normalized island size R/R_c, with m as a parameter.

The TDT model may be used to understand the temperature dependence of the growth mode for heteroepitaxial growth. This is based on the interplay of three characteristic lengths, L_n, L_s, and L_α. Here, L_n is the separation between nucleating islands and is an increasing function of temperature. L_s is the separation between steps on the vicinal substrate,* $L_s = h/\tan\theta$, where h is the step height and θ is the angle of the substrate miscut, and does not depend on temperature. L_α is a length that characterizes the diffusion

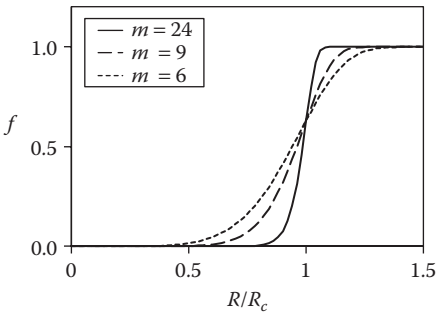


FIGURE 4.13
Fraction of islands exhibiting second-layer nucleation vs. the normalized island size R/R_c, with m as a parameter. R_c is the critical island size. (Adapted from J. Tersoff et al., *Phys. Rev. Lett.*, 72, 266 [1994]. With permission. Copyright 1994, American Physical Society.)

* Even substrates with an exact low-index orientation will typically have a miscut of up to 0.5°, and therefore a finite density of surface steps.

barrier at the island edges and decreases with increasing temperature. With a sufficiently high temperature or step density, $L_n < L_s$; in this case, adatoms can diffuse to the surface steps before nucleating new islands, and the result will be step-flow growth. At a lower temperature, $L_n > L_s$, so the growth mode will be layer by layer (FM), whereby monolayer islands nucleate and then coalesce. At a still lower temperature, $L_\alpha > L_n$, so that multilayer (SK) growth will result.

The case of "reentrant" layer-by-layer growth at still lower temperatures has been attributed to a reduction in the diffusion barrier associated with roughening of the island shapes.[136] TDT offered another explanation for this reentrant behavior. Suppose the islands take on a dendritic shape with arms of characteristic width $2W$, and this width stays roughly constant as the islands grow. Then in this third case of dendritic growth, the critical island size is given by

$$R_{c3} = \left[\left(\frac{2}{\pi} \right) \left(\frac{2D}{F} \right)^{i-1} \frac{L_\alpha^{-i} N_0^{i-3}}{W^{i+2}} \right]^{1/2} . \tag{4.43}$$

Layer-by-layer growth will occur with $R_{c3} > R$, and this can occur at a low temperature if the characteristic width W decreases strongly with temperature.

Figure 4.14 illustrates the expected regimes of growth for various temperatures and substrate miscut angles. If the miscut angle is sufficiently large, the progression from high temperature to lower temperature is as follows: step-flow growth, layer-by-layer, or reentrant Frank-van der Merwe (RFM) growth, multilayer (SK) growth, and finally, reentrant layer-by-layer (RFM) growth. For the miscut angle represented by the dotted horizontal line, the transitions occur at the temperatures T_s, T_α, and T_r, respectively.

The model described here is capable of explaining, at least qualitatively, most of the available experimental evidence in InAs/GaAs(001),[138] $Si_{1-x}Ge_x$/Si(001),[139] and InGaN/GaN(0001).[140]

As in the case of nucleation, kinetic Monte Carlo (KMC) simulations have been applied to predict the growth mode, and specifically surface roughening, in heteroepitaxy.

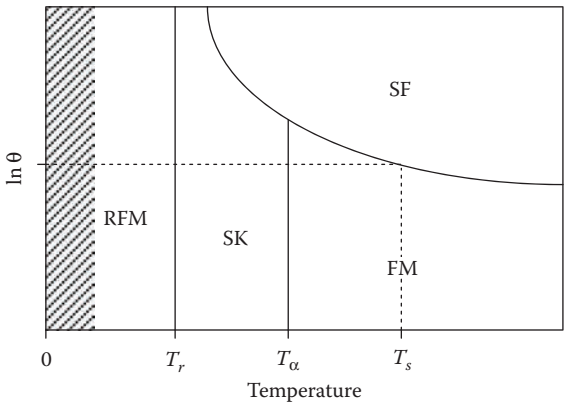


FIGURE 4.14
Regimes of kinetically controlled growth modes, for various values of temperature T and substrate miscut angle θ. (Adapted from J. Tersoff et al., *Phys. Rev. Lett.*, 72, 266 [1994]. With permission. Copyright 1994, American Physical Society.)

Mandreoli et al.[141] have also reported a hybrid approach that combines elements from the rate equation formulation and the KMC approach.

4.5 Low-Temperature Nucleation Layers

Low-temperature buffer layers have been used to minimize the unwanted effects of 3-D nucleation, strain-induced surface roughening, and threading dislocations. In GaAs technology, low-temperature buffer layers have also been studied extensively for their ability to reduce back-gating in metal semiconductor field effect transistors (MESFETs) and MESFET-based integrated circuits,[142,143] but our emphasis here is on the control of morphology, lattice relaxation, and threading dislocations.

4.5.1 Three-Dimensional Growth

In many highly mismatched material systems, the growth mode is 3-D, rather than layer-by-layer or step-flow growth. Therefore, the growth proceeds by the nucleation of a high density of islands, which then coalesce by lateral growth to form a continuous layer. The coalescence of spherical or faceted islands results in a surface with rms roughness r comparable to the separation of the islands, $r \propto d_i \approx 1/\sqrt{\rho_i}$, where ρ_i is the areal density of islands. The density of islands varies exponentially in both the high-temperature and low-temperature regimes, due to the variation of surface mobility and source decomposition, respectively, and peaks at an intermediate value of temperature. Typically, the maximum island density is attained well below the optimum growth temperature for homoepitaxy due to surface mobility considerations. Therefore, the surface smoothness can be greatly improved by growing a thin low-temperature buffer, followed by a thick deposit at a more typical growth temperature. Such a two-step growth process has been used for the growth of GaAs/Si(001),[144] GaAs/Gap (001),[145] InP/Si(001),[146] InP/GaAs(001),[147] ZnTe/GaAs(001),[148] and GaN/Al_2O_3 (0001).[149,150]

Nucleation layers of AlN are commonly used in the heteroepitaxy of nitride semiconductors on sapphire substrates. These AlN buffers avoid the growth of columnar islands and improve the crystal quality of the overgrown GaN. The resulting benefits include an improvement in the electrical and optical properties of the GaN top layer. An additional, but unrelated, benefit of the AlN buffer layer is the compensation of the tensile thermal strain introduced by the sapphire substrate during cooldown. Therefore, the AlN buffer layer can help prevent cracking in thick nitride layers grown on sapphire.

Yoshida et al. first used an AlN buffer layer for the MBE growth of GaN on Al_2O_3 (0001),[149,150] basal-plane sapphire. Here, the AlN buffer was actually grown at a higher temperature (1000°C) than the GaN(700°C). It was found that inclusion of the buffer layer improved the Hall electron mobility by a factor of three, compared with the case of growth directly on sapphire. Moreover, GaN grown with the AlN buffer exhibited up to two orders of magnitude improvement in cathodoluminescence intensity at 360 nm. Amano et al.[151,152] and Koide et al.[153] investigated the structural properties of low-temperature AlN buffer layers used for the growth of GaN on sapphire. They found that the AlN buffer grew as an amorphous layer, thereby suppressing the growth of columnar islands. Further, they observed that heating to the growth temperature for GaN led to the crystallization of the AlN buffer, apparently resulting in a single-crystal surface for heteroepitaxy.

Nakamura[154] applied a GaN low-temperature nucleation layer for the growth of GaN on basal-plane sapphire by MOVPE. The low-temperature GaN nucleation layer was grown at a temperature of 450°C–600°C, whereas the thick top layer of GaN was grown at 1000°C–1030°C. Based on Hall effect measurements of the carrier mobility, the optimum thickness for the nucleation layer was determined to be 20 nm. The GaN grown on an optimized nucleation layer exhibited mirror-smooth morphology over an entire 2 in. wafer. In contrast, GaN grown directly on sapphire without a nucleation layer grew by the coalescence of large hexagonal islands, and exhibited a rough surface.

Kuznia et al.[155] compared the use of GaN and AlN nucleation layers for the MOVPE growth of GaN on sapphire (0001). The nucleation layers used in this study were all grown at 550°C and varied in thickness from 10 nm to 90 nm. The crystallinity of each nucleation layer was investigated by LEED directly after growth and also after a 1 h anneal at 1000°C. It was found that as-grown nucleation layers exhibited no LEED pattern, indicating their amorphous nature. After annealing for 1 h at 1000°C, however, there was a well-defined LEED pattern indicative of a single-crystal layer. Based on electrical measurements (Hall mobility and carrier concentration) made on thick GaN grown upon various nucleation layers, it was found that the optimum nucleation layer thickness was approximately 25 nm (50 nm) for a GaN (AlN) nucleation layer. The use of a nucleation layer greatly improved the crystalline quality (as measured by the x-ray rocking curve full width at half maximum [FWHM]), increased the carrier mobility, and decreased the background doping concentration compared with the case of growth directly on a sapphire substrate. However, either type of nucleation layer (AlN or GaN) gave similar results for all three material parameters. If optimum thickness nucleation layers were used, the AlN nucleation layer gave slightly better results than the GaN nucleation layer.

For the growth of $In_xGa_{1-x}As/GaAs(001)$ with up to 4% mismatch by MOVPE, Ueda et al. showed that 3-D growth could be suppressed by the use of a low-temperature $In_xAl_{1-x}As$ buffer layer grown at 450°C.[156] Whereas growth at 650°C resulted in 3-D growth characterized by faceted islands, low-temperature buffer growth at 450°C yielded a featureless surface that was preserved after annealing and enabled the growth of smooth $In_xGa_{1-x}As$.

ZnO, a wide-bandgap semiconductor of interest for short-wavelength LEDs, is deposited on sapphire or ZnO substrates using a low-temperature nucleation layer. A unique aspect of ZnO epitaxy is that homoepitaxial growth, on a ZnO (0001) substrate, assumes a 3-D growth mode unless the low-temperature nucleation layer is used.[157]

Taylor et al.[158] reported that use of a very low-temperature GaAs buffer layer, grown by migration-enhanced epitaxy at 348K, could suppress VW islanding in GaAs/Si(001), as long as the buffer was deposited to a thickness greater than the "monolithic thickness" h_m. They reported a threading dislocation density of 3×10^6 cm^{-2} and demonstrated p-i-n diodes with ideality of $n = 1.5$ and reverse breakdown electric field of 2.1×10^5 V cm^{-1}.

4.5.2 Surface Roughening

Even in the absence of 3-D growth, there can be strain-induced surface roughening, resulting in corrugations,[159] cross-hatch, or second nucleation of islands on a wetting layer in an SK mode. Here, a reduction in temperature and surface mobility of adatoms may reduce the surface roughening by any of these mechanisms. Low-temperature growth or temperature grading has therefore been used in compositionally graded buffers to improve surface smoothness. Tangring et al.[160] and Song et al.[161] employed low-temperature (400°C and 380°C, respectively) growth of linear $In_xGa_{1-x}As$ buffers on GaAs by MBE. Semenova et al.[162] used low-temperature (400°C) linearly graded $In_xAl_{1-x}As$ for the realization of HEMTs

and quantum dot lasers on GaAs. Wakita et al.[163] used a low-temperature linearly graded $In_xAl_yGa_{1-x-y}As$ buffer in an $In_xGa_{1-x}As$ HEMT on a GaAs substrate. Haupt et al.[164] measured the rms roughness and two-dimensional electron gas (2DEG) mobility for $In_{0.53}Ga_{0.47}As/In_{0.53}Al_{0.47}As$ HEMT structures grown on GaAs with linear and step-graded $In_xAl_{1-x}As$ buffers, for different buffer growth temperatures. Reduced buffer growth temperature, in the range of 300°C–400°C, resulted in reduced roughness and improved 77K mobility.

In order to attain improved surface smoothness of Ge on Si(001), Choi et al.[165] employed a repeated cycle of low-temperature Ge growth, high-temperature Ge growth, and high-temperature hydrogen annealing. Using three such cycles, they obtained rms surface roughness of 0.4–0.6 nm in a 10 μm × 10 μm area.

4.5.3 Lattice Relaxation

The inclusion of a low-temperature buffer layer in the two-step growth process can help relax strain in highly lattice-mismatched material systems. Nomura et al.[145] studied the structural properties of a low-temperature GaAs buffer grown on GaP (001) by MBE. They found that the 25 nm buffer grown at 125°C was amorphous and crystallized by solid phase epitaxy when the temperature was elevated for growth of the high-temperature layer. They found complete relaxation of the mismatch strain during the solid phase epitaxial process.

4.5.4 Threading Dislocations

Reduction of the growth temperature, and therefore surface mobility, can give rise to a large, nonequilibrium density of point defects that are thought to alter the nucleation and dynamics of dislocations.[166]

Shang et al.[167] found that reduced temperature growth of a step-graded $In_xAl_{1-x}As$ buffer layer prevented phase separation and clustering of InAs and AlAs. Czaban et al.[166] found that insertion of a low-temperature (225°C) InP buffer prior to growth of a uniform $InAs_xP_{1-x}$ layer on a step-graded $InAs_yP_{1-y}$ buffer layer improved the x-ray rocking curve FWHM, decreased the crystallographic tilt, and increased the room temperature photoluminescence intensity.

4.6 Surfactants in Heteroepitaxy

Surfactants, or surface-segregated impurities, have a number of applications in heteroepitaxy and engineered heterostructures. The nucleation and growth mode can be modified by the presence of a surfactant.[168] Surfactants may also change the surface reconstruction[169,170] or the misfit dislocation structure in partially relaxed heteroepitaxial layers.[171] In the growth of $In_{0.5}Ga_{0.5}P$, the introduction of a surfactant can suppress the CuPt ordering, which normally occurs in this alloy.[172]

4.6.1 Surfactants and Growth Mode

Surfactants may alter the growth mode for heteroepitaxy by modification of the surface energies for either the substrate or epitaxial layer, if the growth mode is determined

by thermodynamics. Alternatively, a surfactant may change the surface diffusivities or energy barrier at the edge of the islands, if the growth mode is determined by kinetics.

Along the line of thermodynamic considerations, the nucleation and growth mode for heteroepitaxy is determined by $\Delta\gamma$, the areal change in free energy associated with covering the substrate with the epitaxial layer. If γ_e and γ_s are the surface free energies of the epitaxial layer and substrate, respectively, and γ_i is the interfacial free energy for the epitaxial–substrate interface, then

$$\Delta\gamma = \gamma_e + \gamma_i - \gamma_s. \tag{4.44}$$

Often, γ_i may be neglected so that the growth mode will be 2-D (FM) if $\gamma_e < \gamma_s$ and the deposit wets the substrate. On the other hand, a 3-D (VW) mode will result if the deposit does not wet the substrate ($\gamma_e > \gamma_s$). Even if the deposit wets the substrate, the presence of lattice mismatch strain is expected to result in an SK growth mode (layer-by-layer growth followed by islanding) because the creation of islands will partially relieve the mismatch strain.

This has important implications for the growth of heterostructures involving two semiconductors A and B. If A wets B, then B will not wet A, and vice versa. Therefore, in an ABA heterostructure or an ABAB… superlattice, one of the two materials will grow in a 3-D mode, causing a deterioration in the overall structure and its electrical properties.

Copel et al.[168] have proposed that this difficulty may be overcome using a properly chosen surfactant, on the basis of energy considerations. This could be the case if the surfactant atoms have negligible incorporation in a growing crystal of either A or B, and if the surfactant atoms satisfy dangling bonds and reduce the surface energy of crystal A or crystal B. Then the surfactant atoms will "float" on the surface during epitaxy, and may suppress islanding due to surface energy considerations.

On the other hand, surfactants can also modify the growth mode by changing kinetic factors, such as the surface diffusivity for adatoms and the energy barrier for adatoms hopping off the edge of an island (the Ehrlich–Schwoebel barrier).

The effect of surfactants on the growth mode has been most studied in the heteroepitaxial system Ge/Si. Voigtländer and Zinner[173] studied the surfactant-mediated epitaxy of Ge on Si(111) using Sb as the surfactant. The normal growth mode for this heteroepitaxial combination is SK. However, Voigtländer and Zinner found that the Sb surfactant could suppress island formation at a growth temperature of 600°C. However, for growth temperatures of >620°C, the Sb was ineffective in suppressing islanding. They invoked kinetic considerations to explain this result, whereby the Sb surfactant suppresses the surface diffusion of the Ge, thus suppressing island formation if the temperature is sufficiently low. Other studies of surfactant-mediated growth of Ge on Si have shown that group V and VI surfactants decrease the surface diffusion and suppress islanding. However, group III and IV surfactants enhance the surface mobility and have the opposite effect. Hibino et al.[174] studied the surfactant-mediated growth of Ge on Si using Pb. They found that the Pb preabsorption changed the surface structure from Si(111)(7 × 7) to Si(111)$\sqrt{3} \times \sqrt{3}$–Pb. Also, in the growth temperature range 300°C–450°C, the onset of islanding occurred at a thickness of 6 ML without Pb, but occurred at a lower thickness of 4 ML using the Pb surfactant. This result and other work suggests that kinetic considerations are important in determining the growth mode, as well as energetics.

Surfactants have also been investigated as a means of controlling the growth mode in dilute nitride semiconductors such as GaNAs and InGaNAs grown on GaAs(001)

substrates. Tixier et al.[175] studied the use of Bi as a surfactant in the MBE growth of GaNAs and InGaNAs. They found that the Bi suppressed islanding, and step-flow growth could be obtained in $GaN_{0.004}As_{0.996}$ at substrate temperatures as low as 460°C. The Bi also enhanced nitrogen incorporation in the films, though the incorporation of the Bi surfactant was negligible under all conditions studied. Wu et al.[176] studied the use of Sb as a surfactant in this material, also using MBE. The Sb surfactant was found to suppress islanding and improve the photoluminescence intensity of the resulting material. The prevention of islanding was attributed to a reduction of the surface diffusivity for adatoms.[177]

Relatively little work has been reported on surfactant-mediated epitaxy of hexagonal nitride semiconductors on sapphire or SiC substrates. Gupta et al.[178] reported the use of Si as an "antisurfactant" in the MOVPE growth of GaN/AlN/sapphire (0001). Structure growth without the Si surfactant exhibited 2-D growth when grown at 850°C with a V/III ratio of 4.5. Samples grown with the Si surfactant and ramped up to 970°C *after growth* exhibited an island morphology. Widmann et al.[179] and Fong et al.[180] reported the use of In as a surfactant for the growth of GaN/sapphire (0001) by MBE. They found that the use of an In flux during growth promoted 2-D growth and improved the surface roughness and crystallinity of the resulting GaN.

4.6.2 Surfactants and Island Shape

In the VW (3-D) growth of a heteroepitaxial semiconductor, fractal islands are favored at low growth temperature or high incident flux, but compact islands are expected at high temperature or low flux. This behavior has been explained by a diffusion-limited aggregate (DLA) theory that was proposed by Witten and Sander[181] and has been discussed extensively in the literature.[106,182] However, the opposite behavior has been observed in the case of surfactant-mediated growth of Ge/Si(111) using Pb as the surfactant. In this case, fractal islands form at high temperatures, whereas low growth temperatures result in compact islands.[183] Chang et al.[184] explained this behavior by invoking a model of reaction-limited aggregation. In the general case, it appears that surfactants could alter the shapes and sizes of islands by changing diffusion or reaction rates. However, much work remains to clarify the mechanisms and applicability of this approach.

4.6.3 Surfactants and Misfit Dislocations

In a study of the growth of Ge/Si(111), Filimonov et al.[171] found that the use of Bi as a surfactant changed the structure and density of misfit dislocations at the interface. In their study, Ge was grown on Si(111) by MBE at 500°C. In the case of Bi surfactant-mediated epitaxy (Bi-SME), 1 ML of Bi was evaporated from a Knudsen cell prior to epitaxy. The configurations of the interfacial misfit dislocations were inferred from the surface undulations observed in STM micrographs. For the case Bi-SME, the Ge islands exhibited a regular triangular network of misfit dislocations. On the other hand, in the case of conventional growth the misfit dislocations formed a disordered honeycomb network, except near the center of the islands where the triangular network was observed. This difference could come about because of differences in the evolution of the islands. In both cases, the misfit dislocations were found to be 90° Shockley partial dislocations, but the density of misfit dislocations was 25% larger in the case of Bi-SME. The change in the dislocation density could be an *indirect* effect of the surfactant, caused by the suppression of the Si-Ge intermixing. The conventional growth, characterized by greater intermixing, would result in a lower lattice mismatch.

4.6.4 Surfactants and Ordering in InGaP

Several III-V alloys are found to exhibit spontaneous CuPt ordering on the (111) planes when grown by MOVPE.[185] This effect is of practical interest because it alters the bandgap of the material, for a given composition. The pseudobinary semiconductor InGaP ($In_xGa_{1-x}P$ with $x = 0.5$) exhibits a strong tendency for CuPt ordering, with a corresponding change in the bandgap of up to 160 meV.[186] However, the surfactant-mediated epitaxy of this alloy can dramatically suppress the ordering, using either Bi, Sb, or As as the surfactant.[172] This effect has been attributed to the elimination of P dimers on the surface, due to a change in the surface reconstruction. With increasing Sb/Group III ratio, the surface structure changes from $(2 \times n)$ to $\beta 2(2 \times 4)$ and, at a still higher Sb source, flows to a non-(2×4) structure.[169]

The examples described above reveal surfactants to be a powerful tool in modifying the surface structure, growth mode, morphology, and defect structure in heteroepitaxial layers. The field of surfactant-mediated epitaxy is still in its infancy, however, and much theoretical and experimental work remains to be done in this area, especially with the III-nitride materials.

4.7 Quantum Dots and Self-Assembly

Semiconductor quantum dots (QDs) are of great interest for applications including single-electron transistors,[187,188] lasers,[189–191] infrared photodetectors,[192–194] and QD cellular automata (QCA).[195] In all of these applications, the QDs may be fabricated by heteroepitaxial growth in a VW or SK growth mode. In many cases, the resulting dots may have a random distribution on the growth surface, and this is entirely adequate for some device applications. On the other hand, some applications require the precise positioning of QDs, or regular arrays of dots, either one-dimensional (1-D) or 2-D. Self-assembly processes have emerged that appear capable of satisfying these needs, at least to some extent.

The term *self-assembly* has been used extensively in the literature with various meanings. In some cases, the term is used to describe the growth of islands with uniform size, even though their spatial distribution may be quite random. In other cases, *self-assembly* is used to describe the growth of QD islands in a regular geometric pattern, either 1-D or 2-D in nature. (This type of self-assembly has also been called self-organization.)

Practical self-assembly processes serve to alter the surface geometry, chemistry, stress, or perhaps other properties in such a way as to create preferred nucleation sites for islands grown in a VW or SK mode. These processes should therefore be referred to as "guided assembly."

4.7.1 Topographically Guided Assembly of Quantum Dots

Kamins and Williams[196] demonstrated the guided assembly of Ge islands on Si(001), using VPE. In their work, a local oxidation of silicon (LOCOS) process was used to create lines of Si surrounded by silicon dioxide. Some of the Si lines created in this way had submicron width. Next, selective Si epitaxy, using SiH_2Cl_2 and HCl at 850°C and 20 torr, was utilized to produce Si plateaus over the exposed Si lines. These Si plateaus exhibited {311} sidewalls along ⟨110⟩ directions and {110} sidewalls along ⟨100⟩ directions. Ge islands were

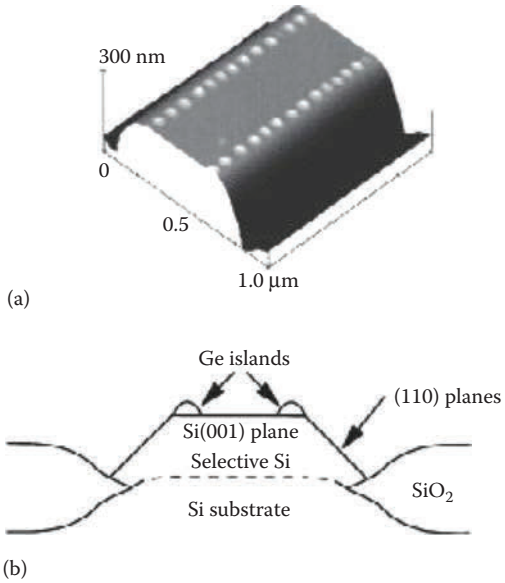


FIGURE 4.15
(a) Three-dimensional AFM micrograph of ordered Ge islands on a Si plateau that was 450 nm wide and had its long axis directed along a ⟨100⟩ direction. The Ge was grown for 120 s at a GeH_4 partial pressure of 2.5×10^{-4} torr. (b) Schematic cross section of the sample. (Reprinted from T. I. Kamins and R. S. Williams, *Appl. Phys. Lett.*, 71, 1201 [1997]. With permission. Copyright 1997, American Institute of Physics.)

next deposited on the Si plateaus using GeH_4 at 600°C and 10 torr. The Ge deposition was carried out either for 60 s at a GeH_4 partial pressure of 5×10^{-4} torr or 120–240 s at a GeH_4 partial pressure of 2.5×10^{-4} torr.

Kamins and Williams found that, for the narrowest silicon lines directed along a ⟨100⟩ direction, the Ge islands grew in two rows near the corners of the plateaus, with an island width of about 75 nm and a regular spacing of 80 nm. Figure 4.15 shows a 3-D atomic force microscopy (AFM) micrograph of ordered Ge islands on a Si plateau that was 450 nm wide and had its long axis directed along a ⟨100⟩ direction. Figure 4.16 shows a 2-D AFM micrograph of the same ordered Ge islands, showing the uniformity of the size and spacing of the islands. This result clearly demonstrates lithographic "demagnification," whereby the self-assembled islands have predictable dimensions and spacings that are much less than the scale of the lithographic features used for their fabrication.

Kamins and Williams also studied Ge island growth on Si lines with different orientations or widths. Wider plateaus exhibited more than two lines of Ge islands, and the ordering of the islands diminished with distance from the plateau edge. Examples are shown in Figure 4.17 for plateaus that were 670–1700 nm wide. The Si plateaus with their long axes oriented along a ⟨110⟩ direction exhibited even less order.

An understanding of the mechanism for this guided assembly technique is of great importance for its application to other geometries or materials. One possible mechanism is based on the kinetics of diffusion for adsorbed species. If there is an Ehrlich–Schwoebel-type energy barrier for the diffusion of Ge adatoms down the sidewall, then reflection of adatoms from this barrier can give rise to enhanced nucleation near the plateau edges. This mechanism would be enhanced for lines along the ⟨100⟩, which exhibit steeper sidewalls. It should be possible to predict the spacing of the islands based on the atomistic nucleation theory.

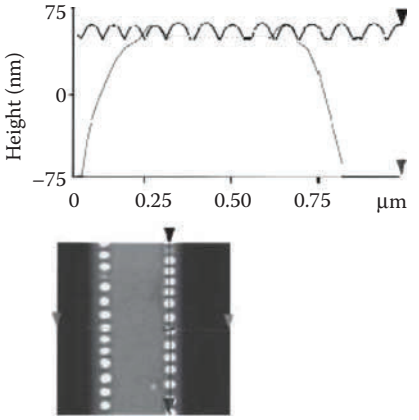


FIGURE 4.16
(a) Two-dimensional AFM micrograph of ordered Ge islands on a Si plateau that was 450 nm wide and had its long axis directed along a ⟨100⟩ direction. The Ge islands grow in a regular pattern with a period of about 80 nm along the ⟨100⟩ direction. The Ge was grown for 120 s at a GeH_4 partial pressure of 2.5×10^{-4} torr. (Reprinted from T. I. Kamins and R. S. Williams, *Appl. Phys. Lett.*, 71, 1201 [1997]. With permission. Copyright 1997, American Institute of Physics.)

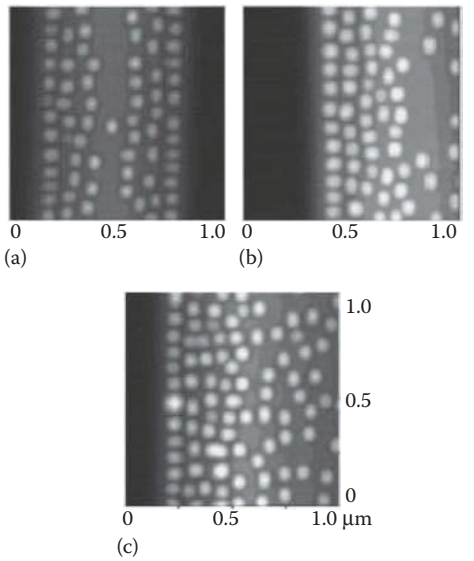


FIGURE 4.17
AFM micrographs of Ge islands on Si plateaus of various widths, which had their long axes directed along a ⟨100⟩ direction. The Ge was grown for 240 s at a GeH_4 partial pressure of 2.5×10^{-4} torr. The plateau width was (a) 670 nm, (b) 1000 nm, and (c) 1700 nm. (Reprinted from T. I. Kamins and R. S. Williams, *Appl. Phys. Lett.*, 71, 1201 (1997). With permission. Copyright 1997, American Institute of Physics.)

Another possible mechanism for the ordering is related to strain relief. According to this explanation, nucleation of islands near the plateau edges is favored because the Si lattice is unconstrained at the sidewall and can distort to reduce the mismatch strain in the islands. This mechanism is also expected to be more effective for the Si lines with the steeper sidewalls, so it is impossible to distinguish between these two mechanisms on this basis.

4.7.2 Stressor-Guided Assembly of Quantum Dots

It has been found that QD nucleation can be strongly influenced by stress in the substrate. In principle, the stress field could be produced by several means. An example of this behavior is the case of Ge QDs grown on a partially relaxed GeSi buffer layer on a Si(001) substrate. Here, a buried array of misfit dislocations exists at the interface between the substrate and the SiGe buffer layer. The Ge dots nucleate preferentially along the lines where the dislocation glide planes intersect the surface of the buffer layer.[197,198] (These are along ⟨110⟩-type directions for the case of SiGe grown on a Si(001) substrate.)

Xie et al.[198] studied the use of a relaxed SiGe layer as a template for fabricating Ge QD arrays. In their work, a relaxed SiGe layer and thin Si cap layer were grown by MBE at 400°C–500°C. Following this, Ge QDs were deposited at 750°C. The AFM micrograph of Figure 4.18 shows the resulting geometry of the Ge QDs after the growth of 1.0 nm Ge coverage (average thickness). The QDs form a rectangular array, with lines of dots parallel to the ⟨110⟩ directions. The positions of the dots correspond closely to the intersections of misfit dislocations at the SiGe/Si interface. This has been attributed to the local strain fields of the dislocations, which reduce the mismatch strain energy in nucleating QDs.

Several aspects of the behavior shown in Figure 4.18 remain incompletely understood at the present time. First, the islands did not organize in this way at lower growth temperatures, though the reason is not clear. Second, the islands observed by Xie et al. exhibited {105} facets when grown at 750°C, even though Mo et al.[199] showed that {105}-facetted Ge huts are not a stable phase at this temperature. Third, the Ge islands position themselves *offset from*, instead of directly over, the places where the dislocations intersect in the relaxed buffer layer. It does not appear that this behavior can be explained on the basis of strain energy alone, but may be controlled in part by the kinetics of surface diffusion.[200] Finally, this method of stressor-guided assembly is ineffective for InAs islands grown on relaxed SiGe buffer layers on Si.[200] While the reason is not clear, it may be related to the difficulty of growing dislocation-free islands of this material, due to the larger lattice mismatch.

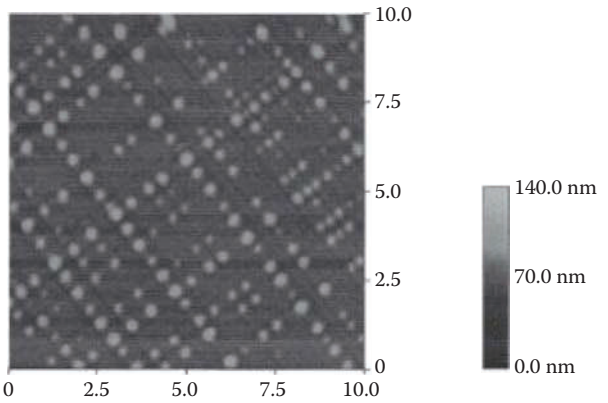


FIGURE 4.18
AFM micrograph showing a regular array of Ge islands on a partially relaxed SiGe layer on a Si(001) substrate. The Ge coverage is 1.0 nm. The Ge QDs nucleate preferentially over the intersections of misfit dislocations in the partially relaxed SiGe layer. (Reprinted from Y. H. Xie et al., *Appl. Phys. Lett.*, 71, 3567 [1997]. With permission. Copyright 1997, American Institute of Physics.)

4.7.3 Vertical Organization of Quantum Dots

Another application of stressor-guided assembly is the fabrication of *vertically assembled* QDs. Here, QDs in a multilayer stack align in vertical columns. In simple terms, the mechanism could be related to the modulation of the stress field by the QDs in one layer, which causes preferential nucleation of QDs in the next layer. The actual detailed mechanism may be much more complex, involving the kinetics of diffusion as well as the stress field.

Xie et al.[201] demonstrated the vertical self-organization of InAs QDs grown on GaAs(001) substrates by MBE. In this work, 2 ML of InAs was deposited on GaAs(001) at 500°C and a grown rate of 0.25 ML s^{-1}. Then a spacer layer was grown, typically consisting of 10 ML GaAs, a 3 ML AlAs marker, and 20 ML GaAs, at 480°C and 0.25 ML s^{-1}. This sequence of layers was grown repeatedly. Figure 4.19 shows a representative cross-sectional TEM micrograph of five sets of vertically organized InAs QDs grown on a GaAs(001) substrate using 36 ML spacer layers.

Mukhametzhanov et al.[202] showed that, in the InAs/GaAs(001) heteroepitaxial system, vertical assembly of InAs QDs could be used to independently manipulate the density and size of QDs. Here, InAs dots were grown at 500°C with a growth rate of 0.22 ML s^{-1}. A GaAs spacer was grown by migration-enhanced epitaxy at 400°C, followed by the growth of another layer of InAs QDs, and finally a GaAs cap. It was shown that the QDs in the second (top) layer aligned with the QDs in the first (bottom) layer. Therefore, the density and size of the QDs could be controlled independently: the deposition time for the first layer of dots controlled the density, and the deposition time for the second layer determined the QD size in that layer.

In the SiGe material system, Teichert et al.[203] demonstrated vertical organization of Ge dots in SiGe/Si multilayer films. Mateeva et al.[204] further studied the vertical organization of Ge dots in SiGe/Si multilayers. Using cross-sectional transmission electron microscopy (TEM) characterization, they showed that the merging of islands of different sizes led to a uniform size distribution after the growth of many periods in these multilayered structures.

4.7.4 Precision Lateral Placement of Quantum Dots

Some device applications require precise placement of QDs rather than the fabrication of dots with uniform size or distribution. In the case of Ge QDs grown on Si(001), this has been achieved by focused-ion-beam micropatterning by Hull et al.[205] and Kammler et al.[206]

In the work of Hull et al. and Kammler et al., clean Si(001) surfaces were irradiated with a Ga^+ focused ion beam, using a beam energy of 25 keV and a beam current of 10 pA (6.2×10^7 ions s^{-1}). AFM images of the surfaces revealed that each irradiated spot contained

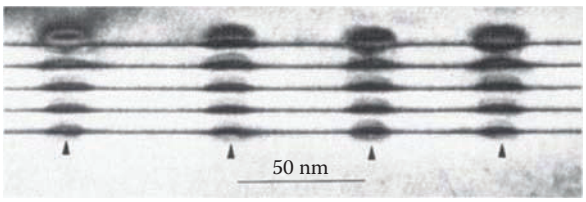


FIGURE 4.19
Cross-sectional TEM micrograph of five sets of vertically organized InAs QDs grown on a GaAs(001) substrate using 36 ML spacer layers. (Reprinted from Q. Xie et al., *Phys. Rev. Lett.*, 75, 2542 (1995). With permission. Copyright 1995, American Physical Society.)

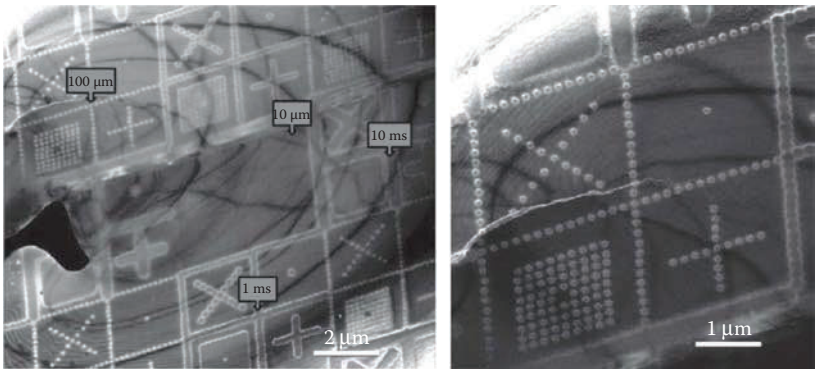


FIGURE 4.20
In situ TEM images of Ge islands on a Si(001) substrate, precisely patterned using a Ga^+ focused ion beam. All patterns were created using a 10 pA Ga^+ beam, but the irradiation times were different for the different regions of the surface as indicated in (a). (b) Micrograph showing an enlargement of the pattern fabricated using a 100 μs irradiation time. (Reprinted from R. Hull et al., *Mater. Sci. Eng. B*, 101, 1 [2003]. With permission. Copyright 2003, Elsevier.)

amorphous material surrounded by a ring of sputtered material. The ring diameter increased from 90 nm for 0.1 ms irradiation to 320 nm for 10 ms irradiation. Following irradiation, the Si(001) was annealed in the range of 600°C to 750°C to recover its crystallinity. Following annealing, Ge QDs were deposited by VPE using digermane at a temperature of 600°C.

Kammler et al. found that for an irradiation time of 0.01 ms the Ge islands formed randomly over the surface and the focused-ion-beam pattern had no influence over their placement. For higher irradiation times (>620 ions/spot), every irradiated spot was occupied by one Ge island, whereas no islands nucleated elsewhere. Figure 4.20 shows Ge QDs that were patterned in this way and demonstrates the remarkable control that is possible. The technique appears to be unaffected by the fill factor or specific pattern to be transferred.

The mechanism underlying this method of precise QD placement is not entirely clear. It could be related to a modification of the surface properties by the implanted Ga, which could diffuse to the surface during the annealing step. Kammler et al. found that the islands formed on the irradiated and annealed areas were smaller and had a larger aspect ratio than those on unirradiated Si, implying a surfactant effect and lending support to this theoretical model. Another possibility, however, is that the implanted Ga ions introduce strain, resulting in stressor-guided assembly. Topography can be ruled out as the mechanism because the irradiated and annealed spots did not develop any topographic relief.

PROBLEMS

1. Sketch the following surface structures: GaAs(001)(2×4), Si(111)(7×7), 6H–SiC(0001) $(\sqrt{3}\times\sqrt{3})R30°$, showing the dimensions in each case.
2. For epitaxial growth of Si at 1000°C, estimate the critical nucleus size for gas phase (homogeneous) nucleation. Assume the equilibrium vapor pressure for Si is ~10^{-3} Pa in order to estimate the supersaturation.

3. Consider the epitaxial growth of $Si_{0.5}Ge_{0.5}$/Si(001) superlattices. Estimate the contact angle for each type of interface. Use Vegard's law to estimate the surface energy of the alloy, and neglect the interfacial energy.
4. For the $Si_{0.5}Ge_{0.5}$/Si superlattices described in Problem 3, describe the expected growth modes for the Si layers and for the $Si_{0.5}Ge_{0.5}$ layers.

References

1. S. Huang, G. Balakrishnan, and D. L. Huffaker, Interfacial misfit array formation for GaSb growth on GaAs, *J. Appl. Phys.*, 105, 103104 (2009).
2. C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chiang, Atomic-scale friction of a tungsten tip on a graphite surface, *Phys. Rev. Lett.*, 59, 1942 (1987).
3. G. Binnig, C. F. Quate, and Ch. Gerber, Atomic force microscope, *Phys. Rev. Lett.*, 56, 930 (1986).
4. S. Kawai, S. Kitamura, D. Kobayashi, and H. Kawakatsu, Dynamic lateral force microscopy with true atomic resolution, *Appl. Phys. Lett.*, 87, 173105 (2005).
5. E. A. Wood, Vocabulary of surface crystallography, *J. Appl. Phys.*, 35, 1306 (1964).
6. Th. Schmidt, R. Kroger, T. Clausen, J. Falta, A. Janzen, M. Kammler, P. Kury, P. Zahl, and M. Horn-von Hoegen, Surfactant-mediated epitaxy of Ge on Si(111): Beyond the surface, *Appl. Phys. Lett.*, 86, 111910 (2005).
7. L. Masson and F. Thibaudau, Role of steps in deposition rate in silane chemical vapor deposition on Si(111), *Phys. Rev. B*, 71, 85314 (2005).
8. J. J. Lander and J. Morrison, Low energy electron diffraction study of silicon surface structures, *J. Chem. Phys.*, 37, 729 (1962).
9. J. A. Appelbaum, G. A. Baraff, and D. R. Hamann, The Si(100) surface. III. Surface reconstruction, *Phys. Rev. B*, 14, 588 (1976).
10. R. E. Schlier and H. E. Farnsworth, Structure and adsorption characteristics of clean surfaces of germanium and silicon, *J. Chem. Phys.*, 30, 917 (1959).
11. J. Levine, Structural and electronic model of negative electron affinity on the Si/Cs/O surface, *Surf. Sci.*, 34, 90 (1973).
12. J. A. Appelbaum, G. A. Baraff, and D. R. Hamann, Si(100) surface reconstruction: Spectroscopic selection of a structural model, *Phys. Rev. Lett.*, 35, 729 (1975).
13. J. A. Appelbaum, G. A. Baraff, and D. R. Haman, The Si(100) surface—Further studies of the pairing model, *Phys. Rev. B*, 15, 2408 (1977).
14. J. C. Phillips, Excitonic instabilities, vacancies, and reconstruction of covalent surfaces, *Surf. Sci.*, 40, 459 (1973).
15. R. A. Wolkow, Direct observation of an increase in buckled dimers on Si(001) at low temperature, *Phys. Rev. Lett.*, 68, 2636 (1992).
16. Y. J. Li, H. Nomura, N. Ozaki, Y. Naitoh, Masami Kageshima, and Y. Sugawara, Origin of p(2 × 1) phase on Si(001) by noncontact atomic force microscopy at 5K, *Phys. Rev. Lett.*, 96, 106104 (2006).
17. T. Nakayama, Y. Tanishiro, and K. Takayanagi, Biatomic layer-high steps on Si(001) 2 × 1 surface, *Jpn. J. Appl. Phys.*, Part 2, 26, L280 (1987).
18. R. Kaplan, LEED study of the stepped surface of vicinal Si(100), *Surf. Sci.*, 93, 145 (1980).
19. B. Z. Olshanetsky and V. I. Mushanov, LEED studies of clean high Miller index surfaces of Si, *Surf. Sci.*, 111, 414 (1981).
20. R. D. Bringans, R. I. G. Uhrberg, M. A. Olmstead, and R. Z. Bachrach, Surface bands for single-domain 2 × 1 reconstructed Si(100) and Si(100):As. Photoemission results for off-axis crystals, *Phys. Rev. B*, 34, 7447 (1986).

21. R. A. Budiman, Interacting dimer rows on terraces: Reconstructed Si(001) surfaces, *Phys. Rev. B*, 72, 35322 (2005).
22. D. Haneman, Electron paramagnetic resonance from clean single-crystal cleavage surfaces of silicon, *Phys. Rev.*, 170, 705 (1968).
23. W. Monch, *Advances in Solid State Physics*, Vol. 13 (Braunschweig: Pergamon-Viewig, 1973), p. 241.
24. M. J. Cardillo, Nature of the Si(111) reconstruction, *Phys. Rev. B*, 23, 4279 (1981).
25. S. Kawai, S. Kitamura, D. Kobayashi, and H. Kawakatsu, Dynamic lateral force microscopy with true atomic resolution, *Appl. Phys. Lett.*, 87, 173105 (2005).
26. J. J. Lander, *Solid State Chemistry*, ed. H. Reiss, Vol. 2 (Oxford: Pergamon, 1965), p. 26.
27. A. J. Van Bommel and F. Meyer, Low energy electron diffraction studies on Ge and Na-covered Ge, *Surf. Sci.*, 6, 57 (1967).
28. M. Taubenblatt, E. So, P. Sih, A. Kahn, and P. Mark, Atomic structure of the annealed Ge(111) surface, *J. Vac. Sci. Technol.*, 15, 1143 (1978).
29. J. J. Lander and J. Morrison, Structures of clean surfaces of germanium and silicon. I, *J. Appl. Phys.*, 34, 1403 (1963).
30. R. Kaplan, Surface structure and composition of β- and 6H-SiC, *Surf. Sci.*, 215, 111 (1989).
31. T.-H. Kim, S. Choi, M. Morse, P. Wu, C. Yi, A. Brown, M. Losurdo, M. M. Giangregorio, and G. Bruno, Impact of unintentional and intentional nitridation of the 6H-SiC$(0001)_{Si}$ substrate on GaN epitaxy, *J. Vac. Sci. Technol. B*, 23, 1181 (2005).
32. X. N. Xie and K. P. Loh, Observation of a 6×6 superstructure on 6H-SiC(0001) by reflection high energy electron diffraction, *Appl. Phys. Lett.*, 77, 3361 (2000).
33. M. Naitoh, J. Takami, S. Nishigaki, and N. Toyama, A $(2\sqrt{3} \times 2\sqrt{13})$ surface phase in the 6H-SiC(0001) surface studied by scanning tunneling microscopy, *Appl. Phys. Lett.*, 75, 650 (1999); *Appl. Phys. Lett.*, 75, 2692 (1999).
34. J. Suda, K. Miura, M. Honaga, Y. Nishi, N. Onojima, and H. Matsunami, Effects of 6H-SiC surface reconstruction on lattice relaxation of AlN buffer layers in molecular beam epitaxial growth of GaN, *Appl. Phys. Lett.*, 81, 5141 (2002).
35. J. D. Hartman, A. M. Roskowski, Z. J. Reitmeier, K. M. Tracy, R. F. Davis, and R. J. Nemanich, Characterization of hydrogen etched 6H-SiC(0001) substrates and subsequently grown AlN, *J. Vac. Sci. Technol. A*, 21, 394 (2003).
36. J. Lu, L. Haworth, D. I. Westwood, and J. E. Macdonald, Initial stages of molecular-beam epitaxy growth of GaN on 6H-SiC(0001), *Appl. Phys. Lett.*, 78, 1080 (2001).
37. K. Jeganathan, M. Shimizu, T. Ide, and H. Okumura, The effect of gallium adsorbate on SiC(0001) surface for GaN by MBE, *Phys. Status Solidi B*, 240, 326 (2003).
38. F. Amy, P. Soukiassian, and C. Brylinski, Atomic cracks and $(2\sqrt{3} \times 2\sqrt{3}) - R\,30°$ reconstruction at 6H-SiC(0001) surface, *Appl. Phys. Lett.*, 85, 926 (2004).
39. V. Ramachandran and R. M. Feenstra, Scanning tunneling spectroscopy of Mott-Hubbard states on the 6H-SiC(0001) $\sqrt{3} \times \sqrt{3}$ surface, *Phys. Rev. Lett.*, 82, 1000 (1999).
40. U. Starke, J. Schardt, J. Bernhardt, M. Franke, K. Reuter, H. Wedler, K. Heinz, J. Furthmuller, P. Kackell, and F. Bechstedt, Novel reconstruction mechanism for dangling-bond minimization: Combined method surface structure determination of SiC(111)–(3×3), *Phys. Rev. Lett.*, 80, 758 (1998).
41. J. Schardt, J. Bernhardt, U. Starke, and K. Heinz, Crystallography of the (3×3) surface reconstruction of 3*C*-SiC(111), 4*H*-SiC(0001), and 6*H*-SiC(0001) surfaces retrieved by low-energy electron diffraction, *Phys. Rev. B*, 62, 10335 (2000).
42. I. Forbeaux, J. M. Themlin, and J. M. Debever, Heteroepitaxial graphite on 6*H* -SiC(0001): Interface formation through conduction-band electronic structure, *Phys. Rev. B*, 58, 16396 (1998).
43. V. M. Bermudez, Structure and properties of cubic silicon carbide (100) surfaces: A review, *Phys. Status Solidi B*, 202, 447 (1997).
44. A. Catellani and G. Galli, Theoretical studies of silicon carbide surfaces, *Prog. Surf. Sci.*, 69, 101 (2002).

45. W. Lu, P. Krüger, and J. Pollmann, Atomic and electronic structure of β-SiC(001)-(3 × 2), *Phys. Rev. B*, 60, 2495 (1999).
46. V. M. Bermudez, Surface science: A metallic semiconductor surface, *Nat. Mater.*, 2, 218 (2003).
47. V. Derycke, P. G. Soukiassian, F. Amy, Y. J. Chabal, M. D. D'Angelo, H. B. Enriquez, and M. G. Silly, Nanochemistry at the atomic scale revealed in hydrogen-induced semiconductor surface metallization, *Nat. Mater.*, 2, 253 (2003).
48. F. Amy and Y. J. Chabal, Interaction of H, O_2, and H_2O with 3C-SiC surfaces, *J. Chem. Phys.*, 119, 6201 (2003).
49. C.-W. Hu, A. Kasuya, S. Suto, A. Wawro, and Y. Nishina, Surface structure of 3C-SiC(111) fabricated by C_{60} precursor: A scanning tunneling microscopy and high-resolution electron energy loss spectroscopy study, *J. Vac. Sci. Technol. B*, 14, 938 (1996).
50. E. S. Hellman, C. D. Brandle, L. F. Schneemeyer, D. Wiesmann, I. Brener, T. Siegrist, G. W. Berkstresser, D. N. E. Buchanan, and E. H. Hartford, $ScAlMgO_4$: An oxide substrate for GaN epitaxy, *MRS Internet J. Nitride Semicond. Res.*, 1, 1 (1996).
51. M. M. Sung, J. Ahn, V. Bykov, J. W. Rabalais, and A. E. Wyckendon, Composition and structure of the GaN $\{000\bar{1}\}-(1\times1)$ surface, *Phys. Rev. B*, 54, 14652 (1996).
52. J. Ahn, M. M. Sung, J. W. Rabalais, D. D. Koleske, and A. E. Wickenden, Surface composition and structure of GaN epilayers on sapphire, *J. Chem. Phys.*, 107, 9577 (1997).
53. M. A. Khan, J. N. Kuznia, D. T. Olsen, and R. Kaplan, Deposition and surface characterization of high quality single crystal GaN layers, *J. Appl. Phys.*, 73, 3108 (1993).
54. W. C. Hughes, W. H. Rowland Jr., M. A. L. Johnson, S. Fujita, J. W. Cook Jr., J. F. Schetzina, J. Ren, and J. A. Edmond, Molecular beam epitaxy growth and properties of GaN films on GaN/SiC substrates, *J. Vac. Sci. Technol. B*, 13, 1571 (1995).
55. M. E. Lin, S. Strite, A. Agarwal, A. Slvador, G. L. Zhaou, N. Teraguchi, A. Rockett, and H. Morkoc, GaN grown on hydrogen plasma cleaned 6H-SiC substrates, *Appl. Phys. Lett.*, 62, 702 (1993).
56. K. Iwata, H. Asahi, S. J. Yu, K. Asami, K. Fujita, M. Fushida, and S. Gonda, High quality GaN growth on (0001) sapphire by ion-removed electron cyclotron resonance molecular beam epitaxy and first observation of (2 × 2) and (4 × 4) reflection high energy electron diffraction patterns, *Ipn. J. Appl. Phys.*, 35, L289 (1996).
57. P. Hacke, G. Feuillet, H. Okumura, and S. Yoshida, Monitoring surface stoichiometry with the (2 × 2) reconstruction during growth of hexagonal-phase GaN by molecular beam epitaxy, *Appl. Phys. Lett.*, 69, 2507 (1996).
58. W. S. Wong, N. Y. Li, H. K. Dong, F. Deng, S. S. lau, C. W. Tu, J. Hays, S. Bidnyk, and J. J. Song, Growth of GaN by gas-source molecular beam epitaxy by ammonia and by plasma generated nitrogen radicals, *J. Cryst. Growth*, 164, 159 (1996).
59. R. J. Molnar, R. Singh, and T. D. Moustakas, Operation of a compact electron cyclotron resonance source for the growth of gallium nitride by molecular beam epitaxy (ECR-MBE), *J. Electron. Mater.*, 24, 275 (1995).
60. A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer, and J. E. Northrop, Determination of wurtzite GaN lattice polarity based on surface reconstruction, *Appl. Phys. Lett.*, 72, 2114 (1998).
61. G. Feuillet, H. Hamaguchi, K. Ohta, P. Hacke, H. Okumura, and S. Yoshida, Arsenic mediated reconstructions on cubic (001) GaN, *Appl. Phys. Lett.*, 70, 1025 (1997).
62. J. R. Arthur and J. J. LePore, GaAs, GaP, and $GaAs_xP_{1-x}$ epitaxial films grown by molecular beam deposition, *J. Vac. Sci. Technol.*, 6, 545 (1969).
63. A. Y. Cho, GaAs epitaxy by a molecular beam method: Observations of surface structure on the (001) face, *J. Appl. Phys.*, 42, 2074 (1971).
64. J. E. Northrup and S. Froyen, Structure of GaAs(001) surfaces: The role of electrostatic interactions, *Phys. Rev. B*, 50, 2015 (1994).
65. T. Hashizume, Q.-K. Xue, J. Zhou, A. Ichimiya, and T. Sakurai, Structures of As-Rich GaAs(001)-(2 × 4) reconstructions, *Phys. Rev. Lett.*, 73, 2208 (1995).

66. T. Hashizume, Q.-K. Xue, A. Ichimiya, and T. Sakurai, Determination of the surface structures of the GaAs(001)-(2 × 4) As-rich phase, *Phys. Rev. B*, 51, 4200 (1995).
67. H. Yamaguchi and Y. Horikoshi, Surface structure transitions on InAs and GaAs(001) surfaces, *Phys. Rev. B*, 51, 9836 (1995).
68. W. G. Schmidt and F. Bechstedt, Geometry and electronic structure of GaAs(001)(2 × 4) reconstructions, *Phys. Rev. B*, 54, 16742 (1996).
69. Y. Garreau, M. Sauvage-Simkin, N. Jedrecy, R. Pinchaux, and M. B. Veron, Atomic structure and faulted boundaries in the GaAs(001) β(2 × 4) surface as derived from x-ray diffraction and line-shape analysis, *Phys. Rev. B*, 54, 17638 (1996).
70. Q.-K. Xue, T. Hashizume, and T. Sakarai, Scanning tunneling microscopy of III-V compound semiconductor (001) surface, *Prog. Surf. Sci*., 56, 1 (1997).
71. V. P. LaBella, H. Yang, D. W. Bullock, P. M. Thibado, P. Kratzer, and M. Scheffler, Atomic structure of the GaAs(001)-(2 × 4) surface resolved using scanning tunneling microscopy and first-principles theory, *Phys. Rev. Lett*., 83, 2989 (1999).
72. M. Gothelid, Y. Garreau, M. Sauvage-Simkin, R. Pinchaux, A. Cricenti, and G. Le Lay, Atomic structure of the As-rich InAs(100)β2(2 × 4) surface, *Phys. Rev. B*, 59, 15285 (1999).
73. C. Ratsch, W. Barvosa-Carter, F. Grosse, J. H. G. Owen, and J. J. Zinck, Surface reconstructions for InAs(001) studied with density-functional theory and STM, *Phys. Rev. B*, 62, R7719 (2000).
74. W. G. Schmidt, S. Mirbt, and F. Bechstedt, Surface phase diagram of (2 × 4) and (4 × 2) reconstructions of GaAs(001), *Phys. Rev. B*, 62, 8087 (2000).
75. R. H. Miwa and G. P. Srivastava, Structure and electronic states of InAs(001)-(2 × 4) surfaces, *Phys. Rev. B*, 62, 15778 (2000).
76. F. Grosse, W. Barvosa-Carter, J. J. Zinck, and M. F. Gyure, Microscopic mechanisms of surface phase transitions on InAs(001), *Phys. Rev. B*, 66, 75321 (2002).
77. Yu. G. Galitsyn, D. V. Dmitriev, V. G. Mansurov, S. P. Moshchenko, and A. I. Toropov, Critical phenomena in the β − (2 × 4) → α − (2 × 4) reconstruction transition on the (001) GaAs surface, *J. Exp. Theor. Phys. Lett*., 81, 629 (2005) (transl. from Russian).
78. L. Li, B.-K. Han, C. H. Li, and R. F. Hicks, Example of a compound semiconductor surface that mimics silicon: The InP(001)-(2 × 1) reconstruction, *Phys. Rev. Lett*., 82, 1879 (1999).
79. L. Li, Q. Fu, B.-K. Han, C. H. Li, and R. F. Hicks, Determination of InP(001) surface reconstruction by STM and infrared spectroscopy of adsorbed hydrogen, *Phys. Rev. B*, 61, 10223 (2000).
80. W. G. Schmidt, P. H. Hahn, F. Bechstedt, N. Esser, P. Vogt, A. Wange, and W. Richter, Atomic structure of InP(001)-(2 × 4): A dimer reconstruction, *Phys. Rev. Lett*., 90, 126101 (2003).
81. G. Chen, S. F. Cheng, D. J. Tobin, L. Li, K. Raghavachari, and R. F. Hicks, Indium phosphide (001)−(2 × 1): Direct evidence for a hydrogen-stabilized surface reconstruction, *Phys. Rev. B*, 68, 121303 (2003).
82. R. R. Wixon, N. A. Modine, G. B. Stringfellow, Theory of surfactant (Sb) induced reconstructions on InP(001), *Phys. Rev. B*, 67, 115309 (2003).
83. X.-Q. Shen, T. Ide, S.-H. Cho, M. Shimizu, S. Hara, H. Okumura, S. Sonoda, and S. Shimizu, Optimization of GaN growth with Ga-polarity by referring to surface reconstruction reflection high-energy electron diffraction patterns, *Jpn. J. Appl. Phys*., 40, L23 (2001).
84. W. A. Harrison, E. A. Kraut, J. R. Waldrop, and R. W. Grant, Polar heterojunction interfaces, *Phys. Rev. B*, 18, 4402 (1978).
85. E. A. Kraut, Summary abstract: Dipole driven diffusion across polar heterojunction interfaces, *J. Vac. Sci. Technol. B*, 1, 643 (1983).
86. D. E. Aspnes and J. Ihm, Biatomic steps on (001) silicon surfaces, *Phys. Rev. Lett*., 57, 3054 (1986).
87. S. F. Fang, K. Adomi, S. Iyer, H. Morkoc, H. Zabel, C. Choi, and N. Otsuka, Gallium arsenide and other compound semiconductors on silicon, *J. Appl. Phys*., 68, R31 (1990).
88. V. Lebedev, J. Jinschek, U. Kaiser, B. Schroter, and W. Richter, Epitaxial relationship in the AlN/Si(001) heterosystem, *Appl. Phys. Lett*., 76, 2029 (2000).
89. P. R. Pukite and P. I. Cohen, Multilayer step formation after As absorption on Si(100): Nucleation of GaAs on vicinal Si, *Appl. Phys. Lett*., 50, 1739 (1987).

90. T. Sakamoto and G. Hashiguchi, Si(001)-2 × 1 single-domain structure obtained by High temperature annealing, *Jpn. J. Appl. Phys.*, 25, L78 (1986).
91. M. Kawabe and T. Ueda, Self-annihilation of antiphase boundary in GaAs on Si(100) grown by molecular beam epitaxy, *Jpn. J. Appl. Phys.*, 26, L944 (1987).
92. D. Saloner, J. A. Martin, M. C. Tringides, D. E. Savage, C. E. Aumann, and M. G. Lagally, Determination of terrace size and edge roughness in vicinal Si{100} surfaces by surface-sensitive diffraction, *J. Appl. Phys.*, 61, 2884 (1987).
93. I. Bhat and W.-S. Wang, Growth of (100) oriented CdTe on Si using Ge as a buffer layer, *Appl. Phys. Lett.*, 64, 566 (1994).
94. M. Grundmann, A. Krost, and D. Bimberg, Observation of the first-order phase transition from single to double stepped Si(001) in metalorganic chemical vapor deposition of InP on Si, *J. Vac. Sci. Technol. B*, 9, 2158 (1991).
95. V. Lebedev, J. Jinschek, J. Krausslich, U. Kaiser, B. Schroter, and W. Richter, Hexagonal AlN films grown on nominal and off-axis Si(001) substrates, *J. Cryst. Growth*, 230, 426 (2001).
96. G. Balakrishnan, M. Mehta, M. N. Kutty, P. Patel, A. R. Albrecht, P. Rotella, S. Krishna, L. R. Dawson, and D. L. Huffaker, Monolithically integrated III-Sb CW super-luminal light emitting diodes on non-miscut Si(100) substrates, *Electron. Lett.*, 32, 244 (2007).
97. H. Noge, H. Kano, T. Kato, M. Hashimoto, and I. Igarashi, Molecular beam epitaxial growth of a GaAs layer free from antiphase domains on an exactly (100)-oriented Si substrate preheated at 1000°C, *J. Cryst. Growth*, 83, 431 (1987).
98. S. Tanaka, C. C. Umbach, J. M. Blakely, R. M. Tromp, and M. Mankos, Fabrication of arrays of large step-free regions on Si(001), *Appl. Phys. Lett.*, 69, 1235 (1996).
99. D. Lee, J. M. Blakely, T. W. Schroeder, and J. R. Engstrom, A growth method for creating arrays of atomically flat mesas on silicon, *Appl. Phys. Lett.*, 78, 1349 (2001).
100. H. H. Farrell, M. C. Tamargo, and J. L. de Miguel, Optimal GaAs(100) substrate terminations for heteroepitaxy, *Appl. Phys. Lett.*, 58, 355 (1991).
101. M. C. Tamargo, J. L. de Miguel, D. M. Hwang, and H. H. Farrell, Structural characterization of GaAs/Znse interfaces, *J. Vac. Sci. Technol. B*, 6, 784 (1988).
102. A. A. Chernov, Nucleation and epitaxy, in *Modern Crystallography III: Crystal Growth* (New York: Springer-Verlag, 1984), p. 48.
103. F. C. Frank and J. H. van der Merwe, One-dimensional dislocations. II. Misfitting monolayers and oriented overgrowth, *Proc. R. Soc. Lond. A*, 198, 216 (1949).
104. M. Volmer and A. Weber, Keimbildung in übersättigten Gebilden, *Z. Physik Chem.*, 119, 277 (1926).
105. J. A. Venables, Kinetic basic of nucleation theory, *Phil. Mag.*, 27, 693 (1973).
106. J. A. Venables, G. D. T. Spiller, and M. Hanbücken, Nucleation and growth of thin films, *Rep. Prog. Phys.*, 47, 399 (1984).
107. J. A. Venables, Nucleation calculations in a pair-binding model, *Phys. Rev. B*, 36, 4153 (1987).
108. M. J. Stowell and T. E. Hutchinson, Nucleation kinetics in thin film growth. II. Analytical evaluation of nucleation and growth behavior, *Thin Solid Films*, 8, 41 (1971).
109. M. J. Stowell, Thin film nucleation kinetics, *Phil. Mag.*, 26, 361 (1972).
110. G. D. T. Spiller, P. Akhter, and J. A. Venables, UHV-SEM study of nucleation and growth of Ag/W(110), *Surf. Sci.*, 131, 517 (1983).
111. M. Hanbücken, M. Futamoto, and J. A. Venables, Nucleation, growth, and the intermediate layer in Ag/Si(100) and Ag/Si(111), *Surf. Sci.*, 147, 433 (1984).
112. J. A. Venables and D. J. Ball, Nucleation and growth of rare-gas crystals, *Proc. R. Soc. Lond. A*, 322, 331 (1971).
113. J. Tersoff, A. W. Denier van der Gon, and R. M. Tromp, Critical island size for layer-by-layer growth, *Phys. Rev. Lett.*, 72, 266 (1994).
114. M. Schneider, I. K. Schuller, and A. Rahman, Epitaxial growth of silicon: A molecular-dynamics simulation, *Phys. Rev. B*, 36, 1340 (1987).
115. M. H. Grabow and G. H. Gilmer, Thin film growth modes, wetting and cluster formation, *Surf. Sci.*, 194, 333 (1988).

116. J. D. Weeks and G. H. Gilmer, Dynamics of crystal growth, *Adv. Chem. Phys.*, 40, 157 (1979).
117. C. Ratsch and J. A. Venables, Nucleation theory and the early stages of thin film growth, *J. Vac. Sci. Technol. A*, 21, S96 (2003).
118. E. Bauer and H. Poppa, Recent advances in epitaxy, *Thin Solid Films*, 12, 167 (1972).
119. I. N. Stranski and L. Krastanov, Zur Theorie der orientierten Ausscheidung von Ionenkristallen aufeinander, *Sitzungsbericht Akademie der Wissenschaften Wien, Math.-naturwiss. Kl. IIb*, 146, 797 (1938).
120. S. Guha, H. Munekata, F. K. LeGoues, and L. L. Chang, Growth mode and dislocation distribution in the ZnSe/GaAs(100) system, *Appl. Phys. Lett.*, 60, 3220 (1992).
121. E. G. Bauer, B. W. Dodson, D. J. Ehrlich, L. C. Feldman, C. P. Flynn, M. W. Geis, J. P. Harbison, et al., Fundamental issues in heteroepitaxy—A Department of Energy, Council on Materials Science Panel Report, *J. Mater. Res.*, 5, 852 (1990).
122. I. Daruka and A.-L. Barabási, Dislocation-free island formation in heteroepitaxial growth: A study at equilibrium, *Phys. Rev. Lett.*, 79, 3708 (1997).
123. I. Daruka, Strained island formation in heteroepitaxy, Doctoral dissertation, University of Notre Dame (1999).
124. H. Brune, Growth modes, in *Encyclopedia of Materials: Science and Technology* (Amsterdam: Elsevier, 2001).
125. G. Wulff, Zur frage der geschwindigkeit des wachsturms under auflösung der kristallflächen, *Z. Krist.*, 34, 449 (1901).
126. J. Tersoff, Stress-induced layer-by-layer growth of Ge on Si(100), *Phys. Rev. B*, 43, 9377 (1991).
127. C. Roland and G. H. Gilmer, Growth of germanium films on Si(001) substrates, *Phys. Rev. B*, 47, 16286 (1993).
128. V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, Spontaneous ordering of arrays of coherent strained islands, *Phys. Rev. Lett.*, 75, 2968 (1995).
129. V. I. Marchenko, Theory of the equilibrium shape of crystals, *Sov. Phys. J. Exp. Theor. Phys.*, 54, 605 (1981).
130. W. Ostwald, Surface diffusion and island density, *Z. Phys. Chem.*, 34, 495 (1900).
131. M. Zinke-Allmang, L. C. Feldman, and M. H. Grabow, Clustering on surfaces, *Surf. Sci. Rep.*, 16, 377 (1992).
132. J. Drucker, Coherent islands and microstructural evolution, *Phys. Rev. B*, 48, 18203 (1993).
133. K. Nakajima, T. Ujihara, N. Usami, K. Fujiwara, G. Sazaki, and T. Shishido, Phase diagram of growth mode for the SiGe/Si heterostructure system with misfit dislocations, *J. Cryst. Growth*, 260, 372 (2004).
134. G. Ehrlich and F. G. Hudda, Atomic view of surface self-diffusion: Tungsten on tungsten, *J. Chem. Phys.*, 44, 1039 (1966).
135. R. L. Schwoebel, Step motion on crystal surfaces. II, *J. Appl. Phys.*, 40, 614 (1969).
136. R. Kunkel, B. Poelsema, L. K. Verheij, and G. Comsa, Reentrant layer-by-layer growth during molecular-beam epitaxy of metal-on-metal substrates, *Phys. Rev. Lett.*, 65, 733 (1990).
137. R. Heitz, T. R. Ramachandran, A. Kalburge, Q. Xie, I. Mukhametzhanov, P. Chen, and A. Madhukar, Observation of reentrant 2D to 3D morphology transition in highly strained epitaxy: InAs on GaAs, *Phys. Rev. Lett.*, 78, 4071 (1997).
138. A. Ohtake and M. Ozeki, In situ observation of surface processes in InAs/GaAs(001) heteroepitaxy: The role of As on the growth mode, *Appl. Phys. Lett.*, 78, 431 (2001).
139. J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy, *J. Vac. Sci. Technol. A*, 2, 436 (1984).
140. R. A. Oliver, M. J. Kappers, C. J. Humphreys, and G. A. Briggs, Growth modes in heteroepitaxy of InGaN on GaN, *J. Appl. Phys.*, 97, 13707 (2005).
141. L. Mandreoli, J. Neugeberger, R. Kunert, and E. Schöll, Adatom density kinetic Monte Carlo: A hybrid approach to perform epitaxial growth simulations, *Phys. Rev. B*, 68, 155429 (2003).
142. M. J. Delaney, C. S. Chou, L. E. Larson, J. F. Jensen, D. S. Deakin, A. S. Brown, W. W. Hooper, M. A. Thompson, L. G. McCray, and S. E. Rosenbaum, GaAs MESFET digital integrated circuits fabricated with low temperature buffer technology, *IEEE Cust. Int. Circ. Conf.*, 18.3.1 (1989).

143. M. J. Delaney, C. S. Chou, L. E. Larson, J. F. Jensen, D. S. Deakin, A. S. Brown, W. W. Hooper, M. A. Thompson, L. G. McCray, and S. E. Rosenbaum, Low-temperature buffer GaAs MESFET technology for high-speed integrated circuit applications, *IEEE Electron Dev. Lett.*, 10, 355 (1989).
144. M. Akiyama, Y. Kawarada, T. Ueda, S. Nishi, and K. Kaminishi, Growth of high quality GaAs layers on Si substrates by MOCVD, *J. Cryst. Growth*, 77, 490 (1986).
145. T. Nomura, K. Wakatuki, M. Miyao, and M. Hagino, Characterization of the GaAs buffer layer grown on GaP by MBE for the two-step growth method, *Appl. Surf. Sci.*, 41/42, 512 (1989).
146. H. Y. Lee, T. E. Crumbaker, M. J. Hafich, G. Y. Robinson, M. M. Al-Jassim, and K. M. Jones, Characterization of InP on Si by MBE, *Proc. SPIE*, 1144, 104 (1989).
147. H. Horikawa, Y. Ogawa, Y. Kawai, and M. Sakuta, Heteroepitaxial growth of InP on a GaAs substrate by low-pressure metalorganic vapor phase epitaxy, *Appl. Phys. Lett.*, 53, 397 (1988).
148. Q. Guo, Y. Sueyasu, T. Tanaka, M. Nishio, and J. Cao, Improvement in the quality of ZnTe epilayers grown on GaAs substrates by introducing a low-temperature buffer layer, *Jpn. J. Appl. Phys.*, 48, 080208 (2009).
149. S. Yoshida, S. Misawa, and S. Gonda, Epitaxial growth of GaN/AlN heterostructures, *J. Vac. Sci. Technol. B*, 1, 250 (1983).
150. S. Yoshida, S. Misawa, and S. Gonda, Improvements on the electrical and luminescent properties of reactive molecular beam epitaxially grown GaN films by using AlN-coated sapphire substrates, *Appl. Phys. Lett.*, 42, 427 (1983).
151. H. Amano, N. Sawaki, I. Akasaki, and Y. Toyoda, Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer, *Appl. Phys. Lett.*, 48, 353 (1986).
152. H. Amano, I. Akasaki, K. Hiramatsu, N. Koide, and N. Sawaki, Effects of the buffer layer in metalorganic vapour phase epitaxy of GaN on sapphire substrate, *Thin Solid Films*, 163, 415 (1988).
153. Y. Koide, N. Itoh, X. Itoh, N. Sawaki, and I. Akasaki, Effect of AlN buffer layer on AlGaN/α-Al_2O_3 heteroepitaxial growth by metalorganic vapor phase epitaxy, *Jpn. J. Appl. Phys.*, 27, 1156 (1988).
154. S. Nakamura, GaN growth using GaN buffer layer, *Jpn. J. Appl. Phys.*, 30, L1705 (1991).
155. N. Kuznia, M. A. Khan, D. T. Olsen, R. Kaplan, and J. Freitas, Influence of buffer layers on the deposition of high quality single crystal GaN over sapphire substrates, *J. Appl. Phys.*, 73, 4700 (1993).
156. T. Ueda, S. Onozawa, M. Akiyama, and M. Sakuta, Heteroepitaxy of InGaAs on GaAs substrate with InAlAs intermediate layer, *J. Cryst. Growth*, 93, 517 (1988).
157. S. H. Park, T. Minegishi, H. J. Lee, D. C. Oh, H. J. Ko, J. H. Chang, and T. Yao, Growth mechanism of ZnO low-temperature homoepitaxy, *J. Appl. Phys.*, 110, 053520 (2011).
158. P. J. Taylor, W. A. Jesser, J. D. Benson, M. Martinka, J. H. Dinan, J. Bradshaw, M. Lara-Taysing, et al., Optoelectronic device performance on reduced threading dislocation density GaAs/Si, *J. Appl. Phys.*, 89, 4365 (2001).
159. B. Lee, J. H. Baek, J. H. Lee, S. W. Choi, S. D. Jung, W. S. Han, and E. H. Lee, Optical properties of InGaAs linear graded buffer layers on GaAs grown by metalorganic chemical vapor deposition, *Appl. Phys. Lett.*, 68, 2973 (1996).
160. I. Tangring, S. Wang, M. Sadeghi, and A. Larsson, 1.27 μ*m* metamorphic InGaAs quantum well lasers on GaAs substrates, *Electron. Lett.*, 42, 691 (2006).
161. Y. Song, S. Wang, I. Tangring, Z. Lai, and M. Sadeghi, Effects of doping and grading slope on surface and structure of metamorphic InGaAs buffers on GaAs substrates, *J. Appl. Phys.*, 106, 123531 (2009).
162. E. S. Semenova, A. E. Zhukov, S. S. Mikhrin, A. Yu. Egorov, V. A. Odnoblyudov, A. P. Vasil'ev, E. V. Nikitina, et al., Metamorphic growth for application in long-wavelength (1.3–1.55 μ*m*) lasers and MODFET-type structures on GaAs substrates, *Nanotechnology*, 15, S283 (2004).
163. A. S. Wakita, H. Rohdin, V. M. Robbins, N. Moll, C.-Y Su, A. Nagy, and D. P. Basile, Low-noise bias reliability of AlInAs/GaInAs MODFETs with linearly graded low-temperature buffer layers grown on GaAs substrates, in *10th International Conference on Indium Phosphide and Related Materials*, Tsukuba, Japan, TuP-38 (1998).

164. M. Haupt, K. Kohler, P. Ganser, S. Emminger, S. Muller, and W. Rothermund, Growth of high quality $Al_{0.48}in_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures using strain relaxed $Al_xGa_yIn_{1-x-y}As$ buffer layers on GaAs, *Appl. Phys. Lett.*, 69, 412 (1996).
165. D. Choi, Y. Ge, J. S. Harris, J. Cagnon, and S. Stemmer, Low surface roughness and threading dislocation density Ge growth on Si(001), *J. Cryst. Growth*, 310, 4273 (2008).
166. J. A. Czaban, D. A. Thompson, and B. J. Robinson, Improved InAsP metamorphic layers grown on an InP substrate using underlying InP grown at low temperatures, *Semicond. Sci. Technol.*, 22, 408 (2007).
167. X. G. Shang, S. D. Wu, C. Liu, W. X. Wang, L. W. Guo, Q. Huang, and J. M. Zhou, Low temperature step-graded InAlAs/GaAs metamorphic buffer layers grown by molecular beam epitaxy, *J. Phys. D Appl. Phys.*, 39, 1800 (2006).
168. M. Copel, M. C. Reuter, E. Kaxiras, and R. M. Tromp, Surfactants in epitaxial growth, *Phys. Rev. Lett.*, 63, 632 (1989).
169. C. M. Fetzer, R. T. Lee, J. K. Shurtleff, G. B. Stringfellow, S. M. Lee, and T. Y. Seong, The use of a surfactant (Sb) to induce triple period ordering in GaInP, *Appl. Phys. Lett.*, 76, 1440 (2000).
170. G. B. Stringfellow, C. M. Fetzer, R. T. Lee, S. W. Jun, and J. K. Shurtleff, Surfactant effects on ordering in GaInP grown by OMVPE, *Proc. Mater. Res. Soc. Symp.*, 583, 261 (2000).
171. S. N. Filimonov, V. Cherepanov, N. Paul, H. Atsaoka, J. Brona, and B. Voigtländer, Dislocation networks in conventional and surfactant-mediated Ge/Si(111) epitaxy, *Surf. Sci.*, 599, 76 (2005).
172. G. B. Stringfellow, J. K. Shurtleff, R. T. Lee, C. M. Fetzer, and S. W. Jun, Surface processes in OMVPE—the frontiers, *J. Cryst. Growth*, 221, 1 (2000).
173. B. Voigtländer and A. Zinner, Structure of the Stranski-Krastanov layer in surfactant-mediated Sb/Ge/Si(111) epitaxy, *Surf. Sci.*, 292, L775 (1993).
174. H. Hibino, N. Shimizu, K. Sumitomo, Y. Shinoda, T. Nishioka, and T. Ogino, Pb preabsorption facilitates island formation during Ge growth on Si(111), *J. Vac. Sci. Technol. A*, 12, 23 (1994).
175. S. Tixier, M. Adamcyk, E. C. Young, J. H. Schmid, and T. Tiedje, Surfactant enhanced growth of GaNAs and InGaNAs using bismuth, *J. Cryst. Growth*, 251, 439 (2003).
176. D. Wu, Z. Niu, S. Zhang, H. Ni, Z. He, Z. Sun, Q. Han, and R. Wu, The role of Sb in the molecular beam epitaxy growth of 1.30–1.55 μm wavelength GaInNAs/GaAs quantum well with high indium content, *J. Cryst. Growth*, 290, 494 (2006).
177. J. C. Harmand, L. H. Li, G. Patriarche, and L. Travers, GaInAs/GaAs quantum-well growth assisted by Sb surfactant: Toward 1.3 μm emission, *Appl. Phys. Lett.*, 84, 3981 (2004).
178. S. Gupta, H. Kang, M. Strassburg, A. Asghar, M. Kane, W. E. Fenwick, N. Dietz, and I. T. Ferguson, A nucleation study of group III-nitride multifunctional nanostructures, *J. Cryst. Growth*, 287, 596 (2006).
179. F. Widmann, B. Daudin, G. Feuillet, N. Pelekanos, and J. L. Rouvière, Improved quality GaN grown by molecular beam epitaxy using In as surfactant, *Appl. Phys. Lett.*, 73, 2642 (1998).
180. W. K. Fong, C. F. Zhu, B. H. Leung, C. Surya, B. Sundaravel, E. Z. Luo, J. B. Xu, and I. H. Wilson, Characteristics of GaN films grown with indium surfactant by RF-plasma assisted molecular beam epitaxy, *Microelectron. Reliab.*, 42, 1179 (2002).
181. T. A. Witten Jr. and L. M. Sander, Diffusion-limited aggregation, a kinetic critical phenomenon, *Phys. Rev. Lett.*, 47, 1400 (1981).
182. H. Brune, Microscopic view of epitaxial metal growth: Nucleation and aggregation, *Surf. Sci. Rep.*, 31, 121 (1998).
183. I. S. Hwang, T.-C. Chang, and T. T. Tsong, Exchange-barrier effects on nucleation and growth of surfactant-mediated epitaxy, *Phys. Rev. Lett.*, 80, 4229 (1998).
184. T.-C. Chang, I. S. Hwang, and T. T. Tsong, Direct observation of reaction-limited aggregation on semiconductor surfaces, *Phys. Rev. Lett.*, 83, 1191 (1999).
185. G. B. Stringfellow, *Organometallic Vapor Phase Epitaxy: Theory and Practice*, 2nd ed. (Boston: Academic Press, 1999), p. 261.
186. L. C. Su, I. H. Ho, N. Kobayashi, and G. B. Stringfellow, Order/disorder heterostructure in $Ga_{0.5}In_{0.5}P$ with $\Delta E_g = 160$ meV, *J. Cryst. Growth*, 145, 140 (1994).

187. H. Ishikuro and T. Hiramoto, Quantum mechanical effects in the silicon quantum dot in a single-electron transistor, *Appl. Phys. Lett.*, 71, 3691 (1997).
188. M. Saitoh, T. Saito, T. Inukai, and T. Hiramoto, Transport spectroscopy of the ultrasmall silicon quantum dot in a single-electron transistor, *Appl. Phys. Lett.*, 76, 1440 (2000).
189. L. Harris, D. J. Mowbray, M. S. Skolnick, M. Hopkinson, and G. Hill, Emission spectra and mode structure of InAs/GaAs self-organized quantum dot lasers, *Appl. Phys. Lett.*, 73, 969 (1998).
190. A. Patanè, A. Polimeni, M. Henini, L. Eaves, P. C. Eaves, P. C. Main, and G. Hill, Thermal effects in quantum dot lasers, *J. Appl. Phys.*, 85, 625 (1999).
191. O. B. Shchekin, G. Park, D. L. Huffaker, and D. G. Deppe, Discrete energy level separation and the threshold temperature dependence of quantum dot lasers, *Appl. Phys. Lett.*, 77, 466 (2000).
192. D. Pan, E. Towe, and S. Kennerly, Normal-incidence intersubband (In,Ga)As/GaAs quantum dot infrared photodetectors, *Appl. Phys. Lett.*, 73, 1937 (1998).
193. D. Pan, E. Towe, and S. Kennerly, A five-period normal incidence (In,Ga)As/GaAs quantum-dot infrared photodetector, *Appl. Phys. Lett.*, 75, 2719 (1999).
194. Z. Chen, O. Baklenov, E. T. Kim, I. Mukhametzhanov, J. Tie, A. Madhukar, Z. Ye, and J. C. Campbell, Normal incidence InAs/$Al_xGa_{1-x}As$ quantum dot infrared photodetectors with undoped active region, *J. Appl. Phys.*, 89, 4558 (2001).
195. G. Bernstein, C. Bazan, M. Chen, C. S. Lent, J. L. Merz, A. O. Orlov, W. Porod, G. L. Snider, and P. D. Tougaw, Practical issues in the realization of quantum-dot cellular automata, *Superlat. Microstruct.*, 20, 447 (1996).
196. T. I. Kamins and R. S. Williams, Lithographic positioning of self-assembled Ge islands on Si(001), *Appl. Phys. Lett.*, 71, 1201 (1997).
197. S. Yu. Shiryaev, F. Jensen, J. L. Hansen, J. W. Petersen, and A. N. Larsen, Nanoscale structuring by misfit dislocations in $Si_{1-x}Ge_x$/Si epitaxial systems, *Phys. Rev. Lett.*, 78, 503 (1997).
198. Y. H. Xie, S. B. Samavedam, M. Bulsara, T. A. Langdo, and E. A. Fitzgerald, Relaxed template for fabricating regularly distributed quantum dot arrays, *Appl. Phys. Lett.*, 71, 3567 (1997).
199. Y. W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Kinetic pathway in Stranski-Krastanov growth of Ge on Si(001), *Phys. Rev. Lett.*, 65, 1020 (1990).
200. Z. M. Zhao, T. S. Yoon, W. Feng, B. Y. Li, J. H. Kim, J. Liu, O. Hulko, et al., The challenges in guided self-assembly of Ge and InAs quantum dots on Si, *Thin Solid Films*, 508, 195 (2006).
201. Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, Vertically self-organized InAs quantum box islands on GaAs(100), *Phys. Rev. Lett.*, 75, 2542 (1995).
202. I. Mukhametzhanov, R. Heitz, J. Zeng, P. Chen, and A. Madhukar, Independent manipulation of density and size of stress-driven self-assembled quantum dots, *Appl. Phys. Lett.*, 73, 1841 (1998).
203. C. Teichert, M. G. Lagally, L. J. Peticolas, J. C. Bean, and J. Tersoff, Stress-induced self-organization of nanoscale structures in SiGe/Si multilayer films, *Phys. Rev. B*, 53, 16334 (1996).
204. E. Mateeva, P. Sutter, J. C. Bean, and M. G. Lagally, Mechanism of organization of three-dimensional islands in SiGe/Si multilayers, *Appl. Phys. Lett.*, 71, 3233 (1997).
205. R. Hull, J. L. Gray, M. Kammler, T. Vandervelde, T. Kobayashi, P. Kumar, T. Pernell, J. C. Bean, J. A. Floro, and F. M. Ross, Precision placement of heteroepitaxial semiconductor quantum dots, *Mater. Sci. Eng. B*, 101, 1 (2003).
206. M. Kammler, R. Hull, M. C. Reuter, and F. M. Ross, Lateral control of self-assembled island nucleation by focused-ion-beam micropatterning, *Appl. Phys. Lett.*, 82, 1093 (2003).

5

Mismatched Heteroepitaxial Growth and Strain Relaxation: I. Uniform Layers

5.1 Introduction

Rarely is heteroepitaxial growth lattice matched. In almost all cases of interest, the epitaxial layer has a relaxed lattice constant that is different from that of the substrate. The lattice mismatch strain* (misfit strain) can be defined as

$$f \equiv \frac{a_s - a_e}{a_e}, \tag{5.1}$$

where:

a_s is the relaxed lattice constant of the substrate
a_e is the relaxed lattice constant of the epitaxial layer

The absolute value of the lattice mismatch may exceed 10%, but is much smaller in many heteroepitaxial systems of practical interest in the field of semiconductors. The mismatch may take on either sign, with some interesting differences observed between tensile ($f>0$) and compressive ($f<0$) systems. This chapter is concerned with several important aspects of mismatched heteroepitaxial growth: (1) the critical layer thickness, (2) the lattice relaxation and introduction of dislocation defects, and (3) the dynamics of dislocation reactions and removal from thick, mismatched layers. Moreover, the main focus of this chapter is on the discussion of equilibrium and kinetically limited relaxation of a single layer with uniform composition on a substrate.

In heteroepitaxial systems with low mismatch ($|f|<1\%|$), the initial growth tends to be coherent, or pseudomorphic. In other words, a thin epitaxial layer takes on the relaxed lattice constant of the substrate within the growth plane. The usual assumption here is that the substrate is thick enough that it is unstrained, although there is evidence that in the case of a thick epitaxial layer such as CdTe/GaAs, the substrate can become strained.[1] Therefore, a pseudomorphic layer exhibits an in-plane strain equal to the lattice mismatch:

$$\varepsilon_{\parallel} = f \text{ (pseudomorphic)}. \tag{5.2}$$

* Two other definitions for the lattice mismatch are often used in the literature; these are $f' \equiv (a_e - a_s)/a_e$ and $f'' \equiv (a_e - a_s)/a_s$. All three definitions yield approximately the same absolute value, but there is a difference in sign that must be accounted for: $f' \approx f'' = -f$.

As the epitaxial layer thickness increases, so does the strain energy stored in the pseudomorphic layer. At some thickness, called the *critical layer thickness* h_c, it becomes energetically favorable for the introduction of *misfit dislocations* in the interface that relax some of the mismatch strain. Beyond the critical layer thickness, therefore, part of the mismatch is accommodated by misfit dislocations (plastic strain) and the balance by elastic strain. In this case,

$$\varepsilon_{\parallel} = f - \delta\,(\text{partially relaxed layer}), \tag{5.3}$$

where δ is the amount of strain relaxation effected by misfit dislocations. A partially relaxed layer is referred to as metamorphic; therefore, the critical layer thickness defines the boundary between pseudomorphic and metamorphic growth. The residual strain is the state of strain present in the heterostructure after relaxation; in a heteroepitaxial layer, the residual strain is generally a function of the lattice mismatch and the layer thickness. It can be calculated based on a thermodynamic model, as long as the growth occurs near thermal equilibrium. In some cases, however, there are kinetic barriers to the lattice relaxation. These are associated with the generation, movement, and interactions of dislocations. Kinetic models have been devised to explain and predict the lattice relaxation behavior in these situations. They predict that the residual strain in the layer will depend on the growth conditions and postgrowth thermal cycling, as well as the mismatch and layer thickness.

In thick, lattice-mismatched heteroepitaxial layers, *most* of the mismatch may be accommodated by misfit dislocations during growth, even if kinetic factors are important. Therefore, the grown layer will be nearly relaxed at the growth temperature. However, the strain measured at room temperature may be quite different if the epitaxial layer and substrate have different thermal expansion coefficients. Then, a thermal strain will be introduced during the cooldown to room temperature. Thermal strain can also be important during thermal cycling either after growth or during growth interruptions.

The introduction of crystal dislocations and other defects is an important aspect of lattice-mismatched heteroepitaxy. The misfit dislocations located at the heterointerface will degrade the performance of any device for which the physical operation depends on this interface. At the same time, any device fabricated in the heteroepitaxial layer will tend to be compromised by the presence of threading dislocations. These are associated with the misfit dislocations and are introduced during the relaxation process. Whereas the misfit dislocations are expected to be present in partially relaxed layers under the condition of thermal equilibrium, threading dislocations are *nonequilibrium defects*. It is therefore possible, at least in principle, to engineer processing approaches to vanish them entirely from a heteroepitaxial layer.

There are important differences between low-mismatch and high-mismatch heteroepitaxial systems, which are not simply a matter of degree. The actual mechanisms of strain relaxation and defect introduction have been found to be different. This is due, at least in part, to the three-dimensional nucleation mode of highly mismatched heteroepitaxial layers. In some high-mismatch heteroepitaxial systems, such as GaSb/GaAs(001), it is found that lattice relaxation occurs in the early stages of growth, by the introduction of pure edge dislocations, rather than by glide of 60° dislocations from the free surface, as often occurs with lower lattice mismatch.

It is often expected that a heteroepitaxial layer will take on the same crystal orientation as its substrate. In practice, both pseudomorphic and partly relaxed layers often exhibit

small misorientations with respect to their substrates. In highly mismatched material systems, gross misorientations are sometimes observed. These come about due to a close match in the atomic spacings for the substrate and the epitaxial layer in different crystallographic directions.

The purpose of this chapter is to explore these issues in more detail, from both the theoretical and experimental perspectives, for uniform composition layers. Chapter 6 extends these concepts to graded layers and multilayered structures, and together these two chapters form the basis for the defect engineering approaches described in Chapter 8, as well as the device applications described in Chapter 9.

5.2 Pseudomorphic Growth and the Critical Layer Thickness

If there is a small lattice mismatch between the epitaxial layer and substrate, and if the growth mode is two-dimensional (layer-by-layer growth), the initial growth will be coherently strained to match the atomic spacings of the substrate in the plane (parallel) of the interface. This situation is depicted schematically in Figure 5.1a for the case of cubic materials.

In Figure 5.1a, the epitaxial layer has a larger lattice constant than the substrate ($a_e > a_s$ and $f < 0$). The substrate is assumed to be sufficiently thick so that it is unstrained by the growth of the epitaxial layer. The unstrained substrate crystal is cubic with a lattice

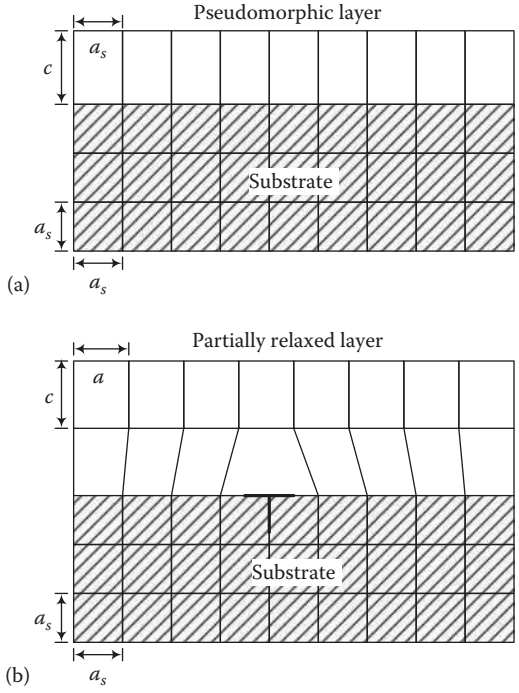


FIGURE 5.1
Growth of a heteroepitaxial layer on a mismatched substrate: (a) pseudomorphic layer and (b) partially relaxed layer with a misfit dislocation.

constant a_s. The pseudomorphic layer matches the substrate lattice constant in the plane of the interface ($a = a_s$) and therefore experiences in-plane biaxial compression. Using the definition for the mismatch adopted here, the in-plane strain is

$$\varepsilon_{\parallel} = f - \delta, \tag{5.4}$$

where δ is the lattice relaxation. In the pseudomorphic layer, for which no lattice relaxation has occurred, $\delta = 0$ and so $\varepsilon_{\parallel} = f$. The epitaxial layer is unconstrained in the direction perpendicular to the interface (the stress in this direction is zero). Therefore, the out-of-plane strain $\varepsilon_{\perp}$ will have the opposite sign compared with $\varepsilon_{\parallel}$ and is given by

$$\varepsilon_{\perp} = -R_B \varepsilon_{\parallel} = -\frac{2C_{12}}{C_{11}} \varepsilon_{\parallel}, \tag{5.5}$$

where R_B is the biaxial relaxation constant of the growing epitaxial layer. The pseudomorphic epitaxial layer is *tetragonally distorted* with an out-of-plane lattice constant c, which is greater than the relaxed lattice constant of the epitaxial layer $c > a_e$.

As the thickness of the growing layer increases, so too does the strain energy in the layer. At some thickness, it becomes energetically favorable for the introduction of misfit dislocations to relax some of the strain. The thickness at which this happens is called the critical layer thickness. For the partially relaxed layer of Figure 5.1b, the in-plane lattice constant of the epitaxial layer has not relaxed to its unstrained value, but it is greater than the substrate lattice constant ($a_e > a > a_s$). So, some of the mismatch is still accommodated by elastic strain. But a portion of the mismatch has been accommodated by misfit dislocations (plastic strain). One such misfit dislocation exists at the interface in Figure 5.1b. Because $a_e > a_s$, this misfit dislocation is associated with an *extra half plane* of atoms in the *substrate*.

As the description above suggests, it is possible to determine the critical layer thickness by the minimization of energy. The total energy is the sum of the strain energy and the energy of the misfit dislocations. We can differentiate the total energy with respect to the strain and determine the minimum (equilibrium) value. The corresponding strain will equal the mismatch at the critical thickness.

Energy calculations for a mismatched epitaxial layer on a substrate were made by Frank and van der Merwe, van der Merwe, and Matthews. None of these models, however, considered the mechanism by which misfit dislocations would be introduced.

The most widely used theoretical model for the critical layer thickness is the force balance model of Matthews and Blakeslee,[2] which is described first. Next, the energy derivation of Frank, van der Merwe, and Matthews is outlined, and it is shown that this derivation gives the same result as the force balance approach, as long as consistent assumptions are made. Finally, the energy derivation of People and Bean will be outlined.

5.2.1 Matthews and Blakeslee Force Balance Model

The Matthews and Blakeslee[2] model is used most often to calculate the critical layer thickness for heteroepitaxy. Here, it is considered that a preexisting threading dislocation in the substrate replicates in the growing epilayer, and can bend over to create a length of misfit dislocation in the interface once the critical layer thickness is reached. This process is shown schematically in Figure 5.2.

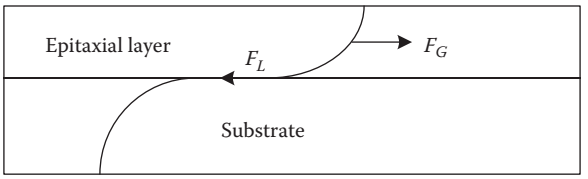


FIGURE 5.2
Bending of a grown-in threading dislocation to create a length of misfit dislocation at the interface between an epitaxial layer and its lattice-mismatched substrate.

For the threading dislocation shown in Figure 5.2, the resolved shear stress acting in the direction of slip is[3]

$$\sigma_{res} = \sigma_{\|} \cos\lambda \cos\phi, \tag{5.6}$$

where:
- $\sigma_{\|}$ is the biaxial stress
- λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface
- ϕ is the angle between the interface and the normal to the slip plane

The glide force acting on the dislocation is

$$F_G = \frac{\sigma_{res} bh}{\cos\phi} = \sigma_{\|} bh \cos\lambda, \tag{5.7}$$

where:
- b is the length of the Burgers vector for the threading dislocation
- h is the film thickness

Assuming biaxial stress in an isotropic and psuedomorphic semiconductor,

$$\sigma_{\|} = \frac{2G(1+\nu)}{(1-\nu)} \varepsilon_{\|} = \frac{2G(1+\nu)}{(1-\nu)} f, \tag{5.8}$$

so that

$$F_G = \frac{2Gbfh(1+\nu)\cos\lambda}{(1-\nu)}, \tag{5.9}$$

where:
- G is the shear modulus
- ν is the Poisson ratio

The line tension of the misfit segment of the dislocation is given by

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{5.10}$$

where:

- G has been assumed to be equal for the epitaxial layer and the substrate
- α is the angle between the Burgers vector and the line vector for the dislocations
- h is the layer thickness

To find the critical layer thickness, we equate the glide force to the line tension for the misfit segment of the dislocation, and solve for the thickness. As a result of this procedure, the critical layer thickness h_c is found to be

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi|f|(1+\nu)\cos\lambda}\left[\ln\left(\frac{h_c}{b}\right)+1\right]. \tag{5.11}$$

For layers with $h < h_c$, the glide force is unable to overcome the line tension, and grown-in dislocations are stable with respect to the proposed mechanism of lattice relaxation. On the other hand, for layers thicker than the critical layer thickness ($h > h_c$), threading dislocations will glide to create misfit dislocations at the interface and relieve the mismatch strain. In the application of Equation 5.11 to (001) zinc blende semiconductors, it is assumed that $\cos\alpha = \cos\lambda = 1/2$ and $b = a/\sqrt{2}$, corresponding to 60° dislocations on $a/2\langle 110\rangle\{111\}$ slip systems. A typical value for the Poisson ratio is $\nu \approx 1/3$. For GaAs, for example, $b = 4.0$ Å and $\nu(001) = 0.312$. Figure 5.3 shows the Matthews and Blakeslee critical layer thickness versus the lattice mismatch strain, calculated assuming $b = 4.0$ Å and $\nu = 1/3$.

5.2.2 Matthews Energy Calculation

To determine the critical layer thickness based on the consideration of energy, we can differentiate the total energy with respect to the strain and determine the minimum (equilibrium) value. The corresponding strain will equal the mismatch strain at the critical thickness.

Matthews[4] derived the critical layer thickness in this manner, starting with the areal strain energy in a pseudomorphic mismatched layer of thickness h, with in-plane strain $\varepsilon_{\|}$ given by

$$E_e = 2G\left(\frac{1+\nu}{1-\nu}\right)h\varepsilon_{\|}^2, \tag{5.12}$$

where:

- G is the shear modulus
- ν is the Poisson ratio

The energy per unit area of a square array of misfit dislocations with average separation S is

$$E_d = \frac{1}{S}\frac{Gb^2\left(1-\nu\cos^2\alpha\right)\left[\ln(R/b)+1\right]}{2\pi(1-\nu)}, \tag{5.13}$$

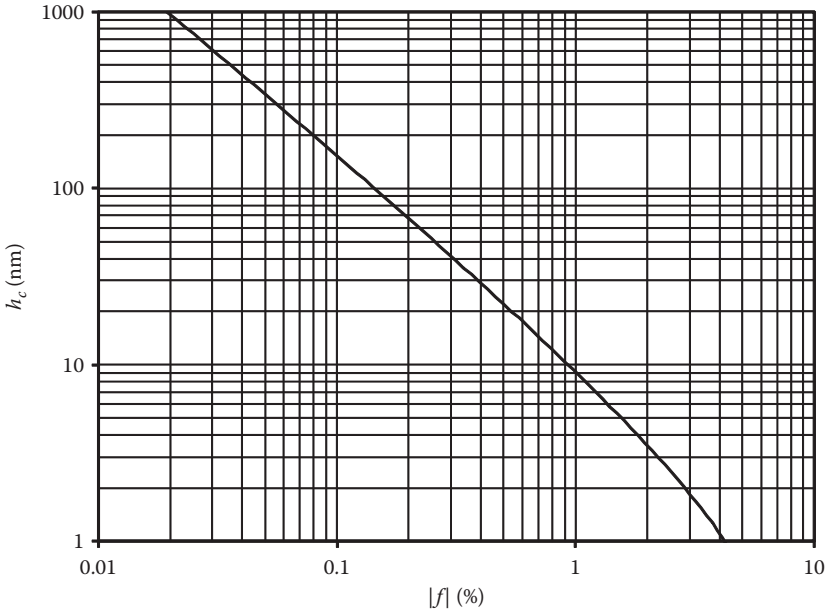


FIGURE 5.3
Matthews and Blakeslee critical layer thickness vs. the lattice mismatch strain, calculated assuming cos α = cos λ = 1/2, b = 4.0 Å, and ν = 1/3.

where:

- α is the angle between the Burgers vector and the line vector for the dislocations
- b is the length of the Burgers vector
- R is the cutoff radius for the determination of the dislocation line energy

This cutoff radius should be taken as the film thickness or the average spacing of the misfit dislocations, whichever is smaller:

$$R = \min(S, h). \tag{5.14}$$

Here, it will be assumed that $R = h$. If the extent of the lattice relaxation is $\delta = f - \varepsilon_{\|}$, then the average spacing of misfit dislocations is

$$S = \frac{b \sin\alpha \sin\phi}{|f - \varepsilon_{\|}|}. \tag{5.15}$$

The total energy of the system is $E_e + E_d$. The condition for energy minimization is

$$\frac{\partial(E_e + E_d)}{\partial \varepsilon_{\|}} = 0. \tag{5.16}$$

Solving, we find the in-plane strain for minimum energy, or the equilibrium strain:

$$\varepsilon_{\parallel,\mathrm{eq}}(h) = \frac{f}{|f|}\frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi h(1+\nu)\sin\alpha\sin\phi}\left[\ln\left(\frac{h}{b}\right)+1\right]. \tag{5.17}$$

Here, the factor $f/|f| = \mathrm{sign}(f)$ accounts for the sign of the strain. The critical layer thickness is the thickness for which $\varepsilon_{\parallel,\mathrm{eq}}(h_c) = f$. Solving,

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi|f|(1+\nu)\sin\alpha\sin\phi}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{5.18}$$

which is exactly the same as the Matthews and Blakeslee critical layer thickness as determined by force balance for a threading dislocation, when taking into account the effective geometrical factors for the effective Burgers vector such that $\sin\alpha\sin\phi = \sin 60^\circ \sin 35.2644^\circ = \cos\lambda = \cos 60^\circ = 1/2$. It should also be noted that often in literature, the expression for h_c is given by

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi|f|(1+\nu)\cos\lambda}\left[\ln\left(\beta\frac{h_c}{b}\right)\right], \tag{5.19}$$

where β is the dislocation core parameter often fitted from experimental results.

The transcendental equations for the critical layer thickness given by Matthews and Blakeslee and by Matthews are typically solved by numerical methods. As an alternative, Braun et al.[5] and Belgacem and Fnaiech[6] proposed a pseudoanalytical approach for the determination of h_c using the Lambert W-function. The Lambert W-function, $W(x)$, is defined by

$$W(x)e^{W(x)} = x, \tag{5.20}$$

where x is the independent variable. To simplify the analysis below, we will use Equation 5.19 for the critical layer thickness and introduce the following definitions:

$$A = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi|f|(1+\nu)\cos\lambda} \tag{5.21}$$

and

$$B = \frac{\beta}{b}. \tag{5.22}$$

Incorporating these two new definitions, the critical layer thickness is simplified to $h_c = A\left[\ln\left(Bh_c\right)\right]$. Using the transformation $h_c \to u = Bh_c$ and by defining $k = -(AB)^{-1}$, we can write the following implicit equation:

$$ku + \ln u = 0, \tag{5.23}$$

which essentially is $-h_c / A + \ln(Bh_c) = 0$. This may be solved using the following assumption:

$$u = \frac{1}{k} W(k) . \tag{5.24}$$

Inserting Equation 5.24 into Equation 5.23 results in

$$k \frac{1}{k} W(k) + \ln\left(\frac{1}{k} W(k) \right) = 0 . \tag{5.25}$$

Taking the antilogarithm and simplifying the equation above yields

$$W(k) e^{W(k)} = k . \tag{5.26}$$

Equation 5.26 is the definition of the Lambert W-function, and therefore the critical layer thickness is given by

$$h_c = -A \cdot W\left(-\frac{1}{AB} \right). \tag{5.27}$$

Substitution of the definitions for A and B yields the following expression for h_c:

$$h_c = -\frac{b\left(1 - \nu \cos^2 \alpha\right)}{8\pi |f| (1+\nu) \cos\lambda} \cdot W\left(-\frac{8\pi |f| (1+\nu) \cos\lambda}{\left(1 - \nu \cos^2 \alpha\right)\beta} \right). \tag{5.28}$$

Because the Lambert W-function is a complex and multivalued function with an infinite number of branches, the exact solution for the critical layer thickness resides in two branches that take real values. Therefore, if x is real, then for the range where $-e^{-1} \le x < 0$, there are two possible real-valued branches of $W(x)$, as shown in Figure 5.4. The branch encompassing the region where $W(x) \le -1$ is referred to as the principal branch $W_0(x)$, whereas the branch where $W(x) \ge -1$ is referred to as $W_{-1}(x)$. Application of Equation 5.28 for the purpose of determining the critical layer thickness requires considerations of the evolution of the Lambert W-function. Because the argument and prefactor of the Lambert W-function are always of the same sign, the $W_{-1}(x)$ branch best describes the dependence of h_c on the lattice mismatch. In addition, it can be seen in Equation 5.28 that $f = 0$ (which corresponds to $x = 0$ in Figure 5.4) must correspond with $h_c \to \infty$ and not $h_c \to 0$. This is apparently only satisfied by the $W_{-1}(x)$ branch. In principle, the Lambert W-function could be applied to graded layers as well as uniform layers, owing to the fact that it is both differentiable and integratable. Thus,

$$\frac{dW}{dx} = \frac{W(x)}{x\left(1 + W(x)\right)} = \frac{e^{-W(x)}}{1 + W(x)}, \text{for } x \ne 0, -\frac{1}{e} , \tag{5.29}$$

where the identity $e^{W(x)} = x / W(x)$ is used and

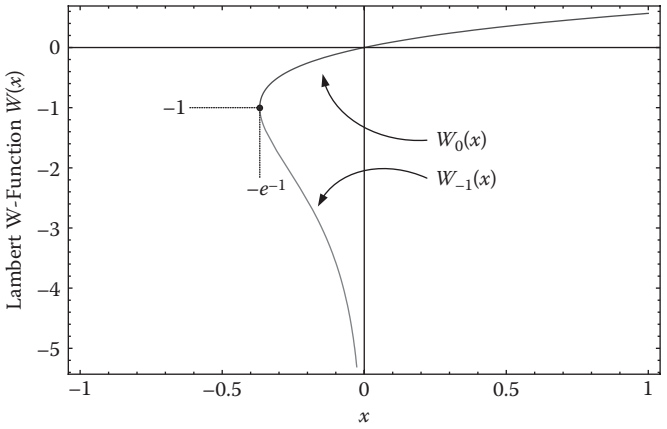


FIGURE 5.4
The real-valued branches $W_0(x)$ and $W_{-1}(x)$ of the Lambert W-function.

$$\int W(x)dx = x\left(W(x)+\frac{1}{W(x)}-1\right)=\left(W(x)^2-W(x)+1\right)e^{W(x)}. \tag{5.30}$$

5.2.3 van der Merwe Model

van der Merwe[7] developed an alternative expression for the critical layer thickness by equating the strain energy in a pseudomorphic film to the interfacial energy of a network of misfit dislocations. The strain energy in the pseudomorphic layer with thickness h was assumed to be

$$E_e = G\left(\frac{1-\nu}{1-2\nu}\right)hf^2, \tag{5.31}$$

where:
- G is the shear modulus
- ν is the Poisson ratio

The areal energy density of a misfit dislocation network was estimated to be

$$E_d \approx \frac{Gb|f|}{4\sqrt{2}\pi(1-\nu)}\ln\left\{\frac{2\pi|f|}{(1-\nu)}\right\}. \tag{5.32}$$

By equating these, van der Merwe found the critical layer thickness to be

$$h_c \approx \frac{(1-2\nu)b}{4\sqrt{2}\pi|f|(1-\nu)^2}\ln\left\{\frac{2\pi|f|}{(1-\nu)}\right\}. \tag{5.33}$$

van der Merwe's predictions are similar to those of Matthews and Blakeslee.

5.2.4 People and Bean Model

People and Bean[8] developed another model for the critical layer thickness by equating the strain energy in a pseudomorphic film to the energy of a dense network of misfit dislocations at the interface. Following Matthews, the strain energy in the pseudomorphic layer with thickness h was assumed to be

$$E_e = 2G\left(\frac{1+\nu}{1-\nu}\right)hf^2, \tag{5.34}$$

where:

G is the shear modulus
ν is the Poisson ratio

People and Bean considered a dense network of misfit dislocations, assumed to have screw character, and with a spacing of $S = 2\sqrt{2}a$. With these assumptions, they calculated the areal energy density of the misfit dislocation array to be

$$E_d \approx \frac{Gb^2}{8\pi\sqrt{2}a}\ln\left(\frac{h}{b}\right). \tag{5.35}$$

Equating this result with the strain energy, and solving for the thickness, they estimated the critical layer thickness to be

$$h_c = \left(\frac{1+\nu}{1-\nu}\right)\left(\frac{1}{16\pi\sqrt{2}}\right)\left(\frac{b^2}{a}\right)\left[\left(\frac{1}{f^2}\right)\ln\left(\frac{h_c}{b}\right)\right], \tag{5.36}$$

where a is the lattice constant for the epitaxial layer. By assuming $a \approx 0.554\,\text{nm}$ and $b \approx 0.4\,\text{nm}$, they obtained

$$h_c = \left(\frac{1.9\times10^{-3}\ \text{nm}}{f^2}\right)\ln\left(\frac{h_c}{0.4\ \text{nm}}\right). \tag{5.37}$$

People and Bean used this expression to calculate the critical layer thickness as a function of composition in $Si_{1-x}Ge_x$/Si(001), for which the lattice mismatch strain is $f = -0.04x$. These results are shown in Figure 5.5, along with the calculations by van der Merwe and by Matthews and Blakeslee. Also shown for comparison are experimental data for several heteroepitaxial material systems. Data for the $Si_{1-x}Ge_x$/Si(001) heteroepitaxial system measured by Bean et al.[9] and Bevk et al.[10] appear to be in agreement with calculations of the People and Bean model. However, the Matthews and Blakeslee model appears to agree with many of the available experimental results. It is known that the effects of finite experimental resolution combined with initially sluggish lattice relaxation can cause experimental results to overestimate the critical layer thickness. This could explain why the People and Bean model is in fair agreement with some experimental results.

The People and Bean model is attractive because its predictions are in fair agreement with some of the experimental results for Si_xGe_{1-x}/Si(001) and also $In_xGa_{1-x}As$/GaAs(001).

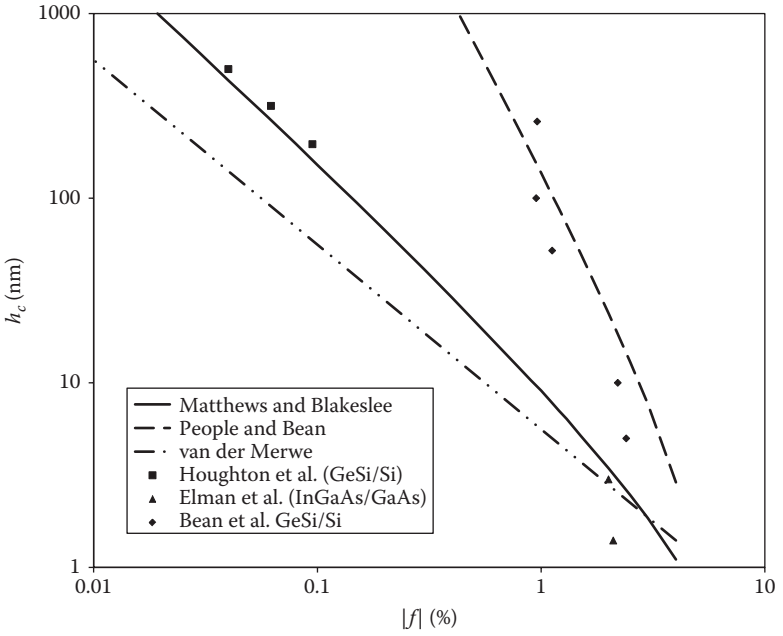


FIGURE 5.5
Critical layer thickness vs. the lattice mismatch strain. The Matthews and Blakeslee[2] critical layer thickness was calculated assuming $\cos\alpha = \cos\lambda = 1/2$, $b = 4.0$ Å, and $\nu = 1/3$. The People and Bean[8] critical layer thickness was calculated using Equation 5.37. The van der Merwe[7] critical layer thickness was calculated using Equation 5.33. The experimental points were measured by Houghton et al.,[11] Elman et al.,[12] and Bean et al.[9]

However, it was developed with the assumption of a dense net of misfit dislocations having a fixed spacing of $2\sqrt{2}a$, and so it is not physical. Since this close spacing of misfit dislocations corresponds to a fully relaxed layer with $|f| \approx 0.062$, the People and Bean model should overestimate the critical layer thickness for heteroepitaxial systems with less than 6.2% mismatch. Moreover, experimental studies of mismatched heteroepitaxial layers have shown that the lattice relaxation occurs gradually with the increase of thickness, and not abruptly.

Although the experimental results shown in Figure 5.5 exhibit considerable scatter, the smallest value for a given mismatch will generally be the most reliable. This is because the combined effects of sluggish lattice relaxation and finite experimental resolution will increase the apparent critical layer thickness obtained by experimentation. The Matthews and Blakeslee model is in fair agreement with the most reliable experimental results, and is the most widely accepted model for the determination of the critical layer thickness.

5.2.5 Effect of the Sign of Mismatch

The models developed by van der Merwe, Matthews, and Blakeslee only consider the absolute value of the lattice mismatch strain, and not its sign. However, it is of technological importance to determine whether the critical layer thickness is different in the tensile and compressive cases. Gendry and Hollinger[13] observed differences in the strain relaxation behavior for compressively and tensile-strained InGaAs on InP, in terms of both the strain versus thickness characteristic and qualitative differences in the surface roughening. They found that in compressive layers, "hump" morphologies such as ridges and islands are

favored, whereas in tensile films, "hollow" morphologies such as grooves, cracks, and pits are favored.

Petruzzello and Leys[14] considered differences in the lattice relaxation mechanisms for compressive and tensile layers arising from the nucleation of Shockley partial dislocations, in diamond and zinc blende semiconductors. (This strain relaxation mechanism is discussed in Section 5.5.4.) However, these differences do not impact the critical layer thickness for the bending over of threading dislocations as considered by Matthews and Blakeslee. On the other hand, Cammarata and Sieradzki[15] modeled the effect of surface tension on the critical layer thickness and showed that, in principle at least, this should make the critical layer thickness smaller for the tensile case and larger for the compressive case. Physically, this asymmetry arises because the surface tension is always compressive. This theoretical treatment will be summarized in what follows.

The elastic strain energy per unit area E_e associated with a uniform elastic strain in an elastically isotropic layer of thickness h with in-plane strain $\varepsilon_{\|}$ is given by[4]

$$E_e = Y\varepsilon_{\|}^2 h, \tag{5.38}$$

where Y is the biaxial modulus, and for an isotropic crystal, $Y = 2G(1+\nu)/(1-\nu)$. The misfit dislocation energy per unit area E_d, for a square array of misfit dislocations along the two ⟨110⟩ directions, and with a spacing such that they relieve an amount of mismatch strain $|\delta| = f - \varepsilon_{\|}$, is given by

$$E_d = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)\left|f-\varepsilon_{\|}\right|}{2\pi(1+\nu)\sin\alpha\sin\phi}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{5.39}$$

where:

- G is the shear modulus
- b is the length of the Burgers vector for the misfit dislocations
- ν is the Poisson ratio
- α is the angle between the Burgers vector and the line vector for the dislocations
- ϕ is the angle between the interface and the normal to the slip plane

The terms E_e and E_d are essentially those used by van der Merwe and Matthews for the calculation of the critical layer thickness by energy balance. Cammarata and Sieradzki introduced another term, E_s, due to the surface energy of the strained heteroepitaxial layer, given by

$$E_s = 2\int \gamma d\varepsilon, \tag{5.40}$$

where γ is the surface energy. It was assumed that γ is isotropic and independent of the strain in the layer, and the critical layer thickness was determined by

$$\frac{\partial\left(E_e + E_d + E_s\right)}{\partial\varepsilon} = 0, \tag{5.41}$$

yielding

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi|f|(1+\nu)\sin\alpha\sin\lambda}\left[\ln\left(\frac{h}{b}\right)+1\right] \pm \frac{\gamma(1-\nu)}{|f|2G(1+\nu)}, \tag{5.42}$$

where the plus and minus signs apply to the compressive and tensile cases, respectively. Apart from the influence of the logarithmic factor, this amounts to the Matthews and Blakeslee critical layer thickness plus or minus a factor proportional to the surface energy.

The variation of the critical thickness with the lattice mismatch strain is plotted in Figure 5.6 for the (a) tensile and (c) compressive cases, assuming a surface energy of $\gamma = 2\,J\,m^{-2}$ and $G = 3\times10^{10}\,Pa$, $\cos\alpha = 1/2, \sin\alpha\sin\phi = 1/2$, $b = 0.4\,\text{nm}$, and $\nu = 1/3$. Using these values,

$$h_c = \frac{0.022\,\text{nm}}{|f|}\left[\ln\left(\frac{h_c}{0.4\,\text{nm}}\right)+1\right] \pm \frac{0.0167\ \text{nm}}{|f|}. \tag{5.43}$$

Also shown is the Matthews and Blakeslee critical layer thickness (b), calculated by neglecting the surface energy. (The second term in Equation 5.31 was neglected.) These results show that, in principle, the surface energy can modify the critical layer thickness, and also create an asymmetry between compressive and tensile films. This work has been extended by Cammarata et al.[16] to include interfacial stresses, for both single heteroepitaxial layers and strained layer superlattices (SLSs). However, the uncertainties inherent in critical layer thickness measurements have hindered experimental verification of this effect.

5.2.6 Critical Layer Thickness in Islands

The theoretical models presented thus far assume that the heterointerface is of "infinite" extent in the lateral directions. Luryi and Suhir[17] showed that in islands with finite lateral size, the critical thickness depends on the island diameter. In their work, Luryi and Suhir calculated the critical layer thickness for mismatched heteroepitaxial islands that

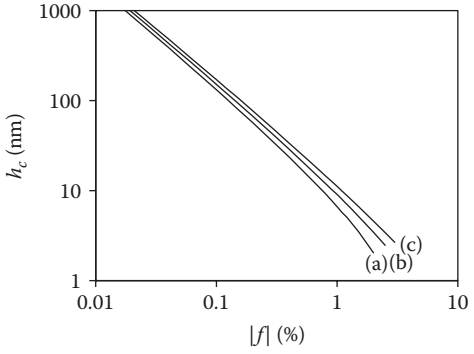


FIGURE 5.6
Critical layer thickness h_c vs. the absolute value of the lattice mismatch strain $|f|$ for (a) tensile films with $\gamma = 2J\,m^{-2}$, (b) tensile or compressive films with $\gamma = 0$ (Matthews and Blakeslee model), and (c) compressive films with $\gamma = 2J\,m^{-2}$. The following values were assumed: $G = 3\times10^{10}\,Pa$, $\cos\alpha = 1/2, \sin\alpha\sin\phi = 1/2$, $b = 0.4\,\text{nm}$, and $\nu = 1/3$.

make rigid contact with the substrate only on round seed pads having a diameter 2*l*. They showed that in a pseudomorphic structure of this sort, the strain in the heteroepitaxial layer decays with distance from the interface. Further, the characteristic length h_e for this decay is on the order of the seed pad dimension. Because of this behavior, the critical layer thickness increases as the seed pads are scaled down in size. For a particular value of the lattice mismatch, there is an island diameter for which the critical layer thickness diverges to infinity, so that structures entirely free from misfit dislocations may be produced.

The analysis of island growth by Luryi and Suhir[17] started with the assumption that the lattice-mismatched heteroepitaxial material makes rigid contact with a noncompliant substrate only at round seed pads having a diameter 2*l*. Here, the *y* axis lies in the plane of the interface, along a major cord of a seed pad. The *z* axis is perpendicular to the substrate and passes through the center of this seed pad. The thickness of the island growth material is *h*. If the substrate is unstrained, then the in-plane stress in the epitaxial deposit is given by

$$\sigma_{\parallel} = f \frac{E}{1-\nu} \chi(y,z) \exp\left(-\frac{\pi z}{2l}\right), \tag{5.44}$$

where:

- f is the lattice mismatch strain
- E is the Young's modulus
- ν is the Poisson ratio, and

$$\chi(y,z) = \begin{cases} 1 - \dfrac{\cosh(ky)}{\cosh(kl)}; & z \le h_e \\ 1; & z \ge h_e \end{cases}, \tag{5.45}$$

where h_e is the effective range for the stress in the *z* direction, to be determined below, and the interfacial compliance parameter *k* is given by

$$k = \left[\frac{3}{2}\left(\frac{1-\nu}{1+\nu}\right)\right]^{1/2} \frac{1}{h_e} \equiv \frac{\zeta}{h_e}. \tag{5.46}$$

The strain energy density per unit volume is

$$u(y,z) = \frac{1-\nu}{E} \sigma_{\parallel}^{\,2}, \tag{5.47}$$

and is maximum at $y = 0$. The strain energy per unit area may be found by integrating over the thickness of the epitaxial deposit and takes on a maximum value at $y = 0$, which is

$$E_s = \int_0^h u(0,z)dz \equiv \frac{E}{1-\nu} f^2 h_e^{\,2}. \tag{5.48}$$

In this calculation, there is little contribution from $z > h_e$, so that it is a good approximation to use the form of $\chi(y,z)$ for $z \le h_e$. The right-hand side of Equation 5.48 defines the characteristic thickness h_e, which is then given implicitly by

$$h_e = h\left\{\left[1-\operatorname{sech}\left(\frac{\zeta l}{h_e}\right)\right]^2\left[1-\exp\left(-\pi h/l\right)\right]\frac{l}{\pi h}\right\} = h\left[\phi\left(\frac{l}{h}\right)\right]^2. \tag{5.49}$$

The right-hand side of this equation defines the "reduction factor" $\phi(l/h)$, which may be determined by numerical solution of Equation 5.49. For $l \gg h$, $h_e \approx h$, but for $l \ll h$,

$$h_e \approx h\left[1-\operatorname{sech}\left(\zeta\pi\right)\right]^2. \tag{5.50}$$

The strain energy per unit area from Equation 5.48 may be used in conjunction with an energy calculation for the critical layer thickness, to find the critical layer thickness h_c^l for an island of radius l. The result is

$$h_c^l = h_c\left[\phi\left(\frac{l}{h_c^l}\right)f\right]. \tag{5.51}$$

In their work, Luryi and Suhir used the People and Bean model for the determination of the critical layer thickness. However, the Matthews energy calculation of the critical layer thickness may also be used, with the result

$$h_c^l = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi\left|\phi\left(l/h_c^l\right)f\right|\left(1+\nu\right)\cos\lambda}\left[\ln\left(\frac{h_c^l}{b}\right)+1\right]. \tag{5.52}$$

The critical thickness is shown as a function of the lattice mismatch, with the island diameter $2l$ as a parameter, in Figure 5.7. The Matthews and Blakeslee curve corresponds to $2l \to \infty$. For nanometer-scale islands, the critical layer thickness can be increased significantly. Also, at a given mismatch, there is a critical island diameter for which the critical thickness diverges to infinity. The critical island size is plotted as a function of the lattice mismatch strain in Figure 5.8.

5.2.7 Critical Layer Thickness in Nitride Semiconductors

There have been a number of attempts to apply a Matthews and Blakeslee–type force balance model to hexagonal nitride semiconductors. For example, Srinivasan et al.[18] applied the force balance model to *c*-plane InGaN/GaN growth, and compared these calculations with experiments for the case of three possible slip systems having $\cos\lambda \neq 0$. They found the best agreement when the $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ slip system was assumed, with $\alpha = 90°$, $\phi = \tan^{-1}(a/c)$, $b = \sqrt{a^2+c^2}$, and $\lambda = 90° - \phi$. On this basis, they concluded that the $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ slip system was active in this material.

On the other hand, the Matthews and Blakeslee force balance model discussed so far was developed for cubic materials with the assumption of isotropic behavior and by making usual assumptions regarding the core energy for dislocations in cubic materials. So although these models serve as a good first approximation when calculating the critical layer thickness in $In_xGa_{1-x}N$/GaN or $Al_xGa_{1-x}N$/GaN, the importance of these materials in light-emitting diodes (LEDs) and high-electron-mobility transistors (HEMTs) has led to the development of refined models by Holec et al.[19,20] and Ye et al.[21]

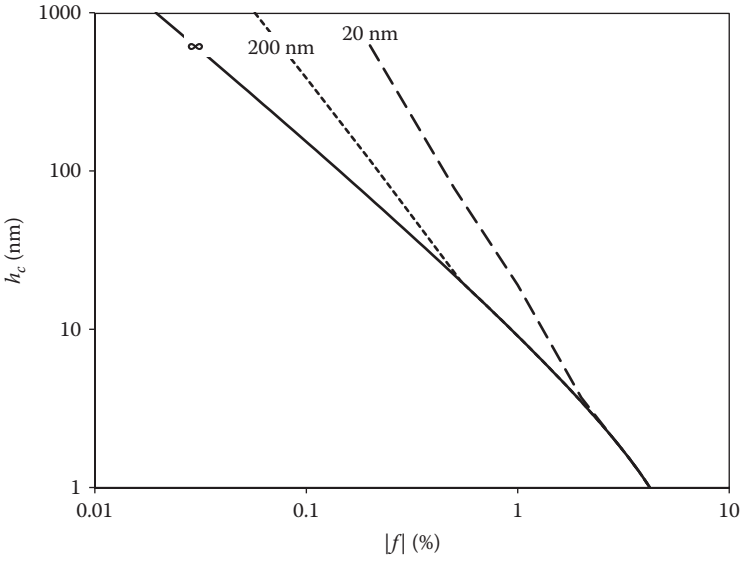


FIGURE 5.7
Critical layer thickness as a function of lattice mismatch strain, with island diameter $2l$ as a parameter. The case of $2l = \infty$ corresponds to planar growth (the Matthews and Blakeslee limit).

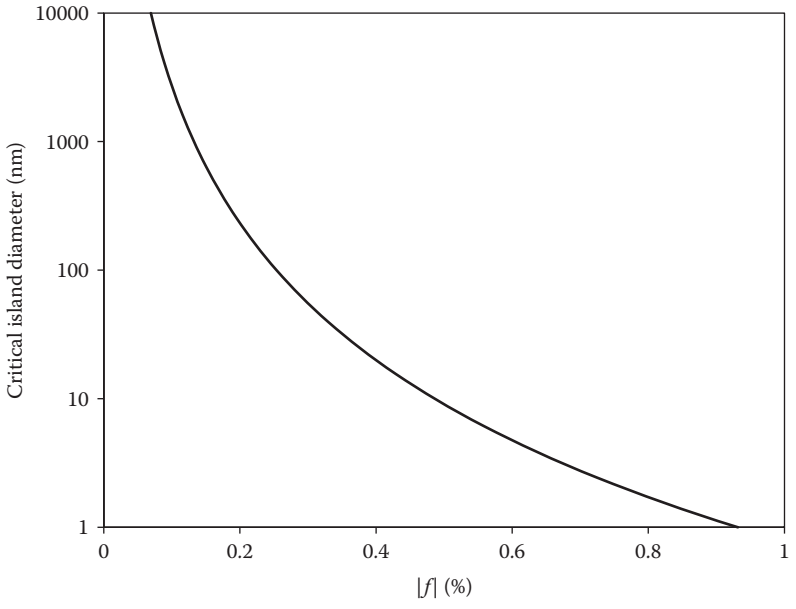


FIGURE 5.8
Critical island diameter as a function of lattice mismatch strain. For islands equal to or less than this diameter, the critical layer thickness diverges to infinity, and pseudomorphic structures may be grown with arbitrary thickness.

The critical layer thickness may be determined by energy minimization,[17] using the criterion

$$E_d = E_e, \tag{5.53}$$

where:

E_d is the dislocation energy per unit length
E_e is the strain energy relieved by the dislocation per unit length

The dislocation energy per unit length includes elastic and core terms and is given by

$$E_d = \left(Ab_{e,\|}^2 + Bb_{e,\perp}^2 + Cb_s^2\right)\ln\left(\frac{R}{r_c}\right) + E_{core}, \tag{5.54}$$

where A, B, and C are material constants related to the elastic stiffness coefficients, and are given in Figure 5.9 for the case of $In_xGa_{1-x}N$; $b_{e,\|}$ and $b_{e,\perp}$ are the in-plane and out-of-plane edge components of the Burgers vector; b_s is the screw component of the Burgers vector; R is the film thickness or dislocation separation, whichever is smaller; and r_c is the cutoff radius for the dislocation core. In some work, it was assumed that $r_c \approx b/2$, but based on the calculations of Belabbas et al.,[22] Holec et al. assumed $r_c \approx b$ for $a+c$-type dislocations, and based on the calculations of Lymperakis[23] and Kioseoglou et al.,[24] Holec et al. assumed $r_c \approx 2b$ for a-type dislocations (Table 5.1).

The strain energy relieved per unit length of dislocation is

$$E_e = \frac{\left(C_{11}+C_{12}\right)C_{33} - 2C_{13}^2}{C_{33}} f b_{e,\|} h. \tag{5.55}$$

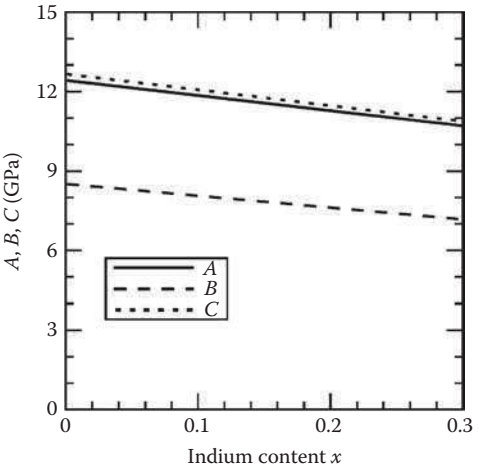


FIGURE 5.9
Values of the coefficients in the dislocation line energy calculation, for $In_xGa_{1-x}N$. (Reprinted from D. Holec et al., *J. Cryst. Growth*, 303, 314 [2007]. With permission. Copyright 2007, Elsevier.)

TABLE 5.1
Parameters of the 5/7-Atom Ring Dislocation Cores in Nitride Semiconductors as Reported by Belabbas et al.,[20] Lymperakis,[21] Kioseoglou et al.,[22] and Holec et al.[18]

Material	Dislocation	R_c (nm)	E_{core} (eV/Å)
GaN	$(a + c)$ type	0.72	3.12
GaN	a type	0.60	1.61
AlN	a type	0.83	1.71
InN	a type	0.54	1.66

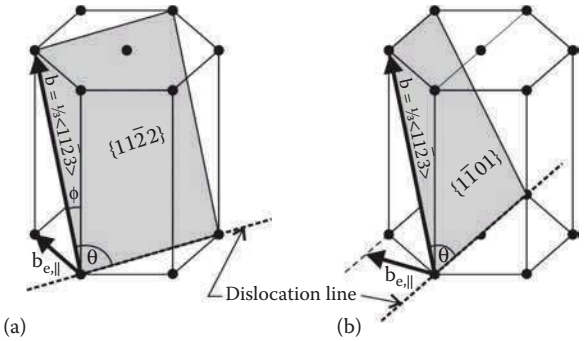


FIGURE 5.10
The two prismatic slip systems considered in the calculation of the critical layer thicknesses for $In_xGa_{1-x}N$/GaN and $Al_xGa_{1-x}N$/GaN by Holec et al.[18] (a) $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ and (b) $1/3\langle 11\bar{2}3\rangle\{1\bar{1}01\}$. (Reprinted from D. Holec et al., *J. Appl. Phys.*, 104, 123514 [2008]. With permission. Copyright 2008, American Institute of Physics.)

In calculating the critical layer thicknesses in $In_xGa_{1-x}N$/GaN and $Al_xGa_{1-x}N$, Holec et al. considered the basal slip system $1/3\langle 11\bar{2}0\rangle\{0001\}$ and two prismatic slip systems, $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ and $1/3\langle 11\bar{2}3\rangle\{1\bar{1}01\}$. The two prismatic slip systems are illustrated in Figure 5.10.

Calculated values of the critical thickness in $Al_xGa_{1-x}N$/GaN are given for each of the three slip systems in Figure 5.11. Theoretical calculations are compared with experimental values for $In_xGa_{1-x}N$/GaN and $Al_xGa_{1-x}N$/GaN in Figures 5.12 and 5.13.

The calculated critical layer thicknesses for $Al_xGa_{1-x}N$/GaN(0001) are in fair agreement with experimental results if the $1/3\langle 11\bar{2}3\rangle\{11\bar{2}2\}$ slip system is considered and the dislocation core energy is included. For the case of $In_xGa_{1-x}N$/GaN(0001), some of the experimental data points for samples absent misfit dislocations (MDs) are considerably thicker than the predicted equilibrium critical thickness. This could be explained by the combined effects of initially slugglish relaxation and finite experimental resolution for the methods employed, as in the case of cubic semiconductors.

Whereas the treatment above considers only the strain energy and line energy contributions, Kioseoglou et al.[36] considered contributions due to the interfacial energy. In that work, they showed that the critical layer thickness for AlN/GaN(0001) depends on the polarity of the interface, with the III-polarity interface exhibiting a larger critical layer thickness than the N-polarity interface.

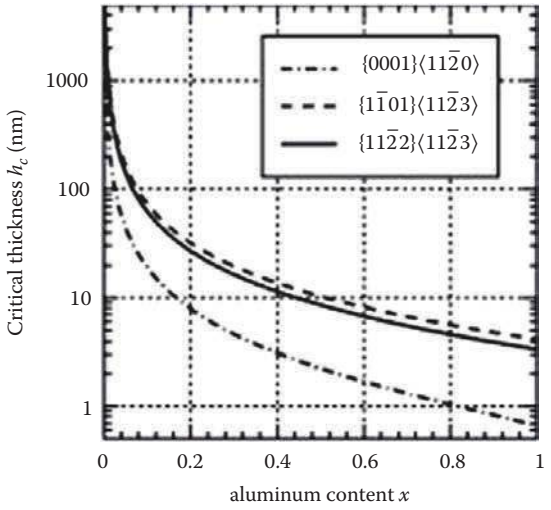


FIGURE 5.11
Theoretical predictions of the critical layer thickness for $Al_xGa_{1-x}N/GaN(0001)$ for the basal slip system 1/3⟨11$\bar{2}$0⟩ {0001} and the prismatic slip systems 1/3⟨11$\bar{2}$3⟩ {1$\bar{1}$01} and 1/3⟨11$\bar{2}$3⟩ {11$\bar{2}$2}. (Reprinted from D. Holec et al., *J. Appl. Phys.*, 104, 123514 [2008]. With permission. Copyright 2008, American Institute of Physics.)

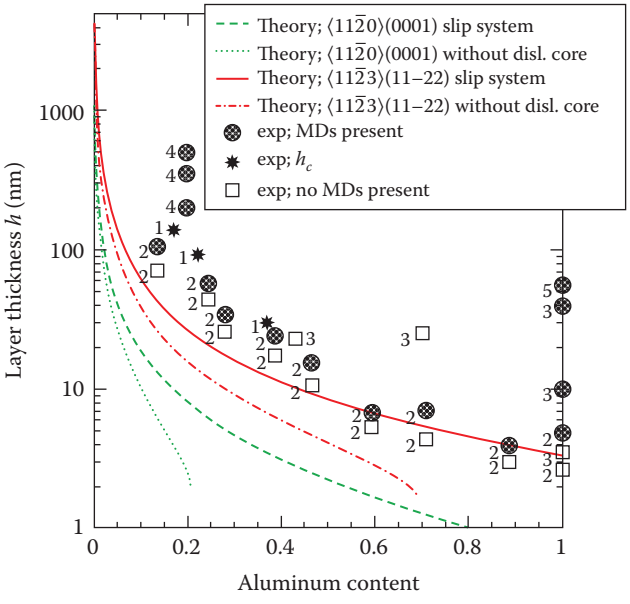


FIGURE 5.12
Comparison of critical layer thickness calculations to experimental data for $AlxGa_{1-x}N/GaN(0001)$. Disl., dislocation. Experimental data were taken from Floro et al.,[25] Lee et al.,[26] Vennegues et al.,[27] Bethoux and Vennegues,[28] and Gherasimova et al.[29] (Reprinted from D. Holec et al., *J. Appl. Phys.*, 104, 123514 [2008]. With permission. Copyright 2008, American Institute of Physics.)

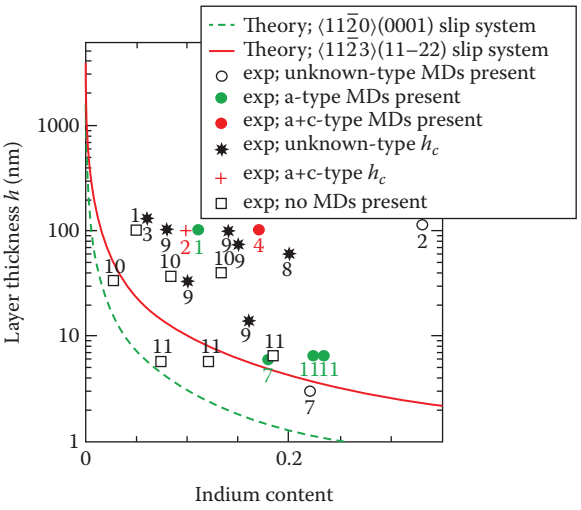


FIGURE 5.13
Comparison of critical layer thickness calculations to experimental data for $In_xGa_{1-x}N$/GaN(0001). Experimental data were taken from Liu et al.,[30] Srinivasan et al.,[16] Lu et al.,[31] Liu et al.,[32] Cho and Yang,[33] Jahnen et al.,[34] Costa et al.,[35] Parker et al.,[36] and Reed et al.[37] (Reprinted from D. Holec et al., *J. Appl. Phys.*, 104, 123514 [2008]. With permission. Copyright 2008, American Insitute of Physics.)

5.3 Dislocation Sources

It is well established that lattice relaxation commences after the critical layer thickness is exceeded in a mismatched heteroepitaxial layer. The strain relaxation occurs by the introduction of misfit dislocations at the interface, but where do these dislocations come from? Matthews and Blakeslee derived the critical layer thickness based on the assumption that substrate dislocations bend over to produce misfit dislocations in the interface. However, heteroepitaxial layers may contain 10^5 times the threading dislocation density of their substrates; thus, substrate dislocations are usually not the sole source. Instead, dislocation nucleation (homogeneous or heterogeneous) or dislocation multiplication must take place during mismatched heteroepitaxy. Heterogeneous nucleation involves an existing defect, whereas homogeneous nucleation does not. For wurtzite semiconductors, lack of high-quality GaN substrates (threading dislocation densities upwards of 10^9 cm^{-2}) often implicates threading dislocations as the primary source of misfit dislocations at the heterointerface. Some studies have revealed cracks or other defects in the film as the sole sources of misfit dislocation nucleation. Srinivasan et al.[18] experimented with ELOG GaN substrates and found that none of the misfit dislocations (MDs) generated were associated with threading dislocations in the substrate. The following sections will review these sources of dislocations, and it will be shown that heterogeneous nucleation and multiplication can both be important.

5.3.1 Homogeneous Nucleation of Dislocations

One mechanism that has been considered for the introduction of misfit dislocations is the homogeneous nucleation of half loops at the surface. The glide of such a half loop to the interface would result in a misfit dislocation segment with two associated threading

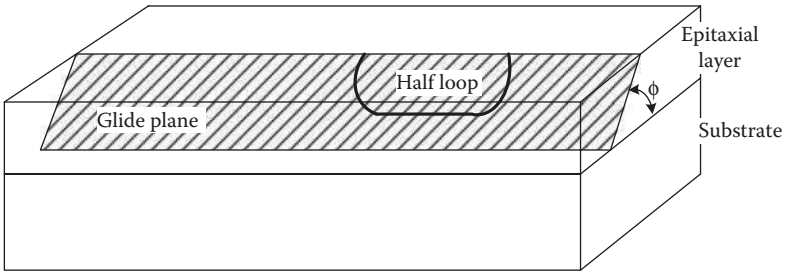


FIGURE 5.14
A dislocation half loop of radius *R* that has nucleated at the surface of a heteroepitaxial layer, lying on its glide plane.

dislocations. The nucleation of such half loops has been considered in detail by Matthews[4,39] and Matthews et al.[40]

Consider a dislocation half loop that has nucleated at the surface of a heteroepitaxial layer as shown in Figure 5.14. This situation is analogous to the nucleation of a deposit on a substrate as described in Section 4.3.2. Thus, there is a critical half-loop radius R_c above which the loop will continue to grow until it reaches the interface, creating a length of misfit dislocation. Subcritical loops will shrink and disappear. The homogeneous nucleation of dislocation half loops will occur only if the thermal energy is sufficient for the spontaneous formation of loops having a radius equal to the critical radius.

The critical radius and the associated half-loop energy were calculated by Matthews as follows. The formation of the dislocation half loop with a radius *R* involves the line energy*

$$E_l = \frac{Gb^2R}{8}\left[\frac{(2-\nu)}{(1-\nu)}\right]\ln\left(\frac{R}{b}+1\right), \tag{5.56}$$

where:

- G is the shear modulus
- b is the length of the Burgers vector
- ν is the Poisson ratio

This is offset by the strain energy released by the formation of the half loop, which is

$$E_e = \frac{\pi RGb(1+\nu)}{(1-\nu)}\cos\lambda\cos\phi, \tag{5.57}$$

where λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface, and ϕ is the angle between the interface and the normal to the slip plane. If the loop is imperfect, there is a stacking fault on the inside of it with energy

$$E_{sf} = \frac{\pi R^2\sigma}{2}, \tag{5.58}$$

* Matthews used a slightly different expression for the self-energy of the dislocation half loop. However, the end result is not changed significantly by this difference.

where σ is the stacking fault energy per unit area. The energy of the surface step created by the introduction of the loop is

$$E_s = 2R\gamma b \sin\alpha\,, \tag{5.59}$$

where:

- γ is the surface energy per unit area
- α is the angle between the Burgers vector and the line vector for the misfit portion (bottom) of the loop

The total energy of the loop is found by summing these terms,

$$E = E_l - E_e + E_{sf} + E_s\,. \tag{5.60}$$

This energy is zero for $R = 0$, increases to a maximum value of E_{crit} at the critical half-loop size R_{crit}, and then decreases for larger values of R. The critical half-loop size can then be found by

$$\frac{\partial E}{\partial R} = 0\,. \tag{5.61}$$

If a complete dislocation is assumed (no stacking fault is involved), the critical radius is given by

$$R_{\text{crit}} = \frac{\left(\frac{Gb^2}{8}\right)(2-\nu)\left[\ln\left(\frac{R}{b}\right)+2\right]+2\sigma(1-\nu)b\sin\alpha}{2\pi G\left(1+\nu\varepsilon b\cos\lambda\cos\phi\right)}\,. \tag{5.62}$$

The activation energy for the nucleation of half loops may be determined from

$$E_{\text{crit}} = E\left(R_{\text{crit}}\right)\,. \tag{5.63}$$

If it is assumed that the available thermal energy does not exceed 50 kT, then the homogeneous nucleation of half loops is not expected to occur unless the mismatch strain is greater than 1.5% (at room temperature).

Due to the large amount of energy involved in this process, it is believed that the homogeneous nucleation of half loops will be insignificant in most heteroepitaxial materials. Instead, it is likely that other processes produce misfit dislocations and start to relieve the misfit strain before this homogeneous nucleation process can become active.

5.3.2 Heterogeneous Nucleation of Dislocations

The calculations of the previous section show that the homogeneous nucleation of dislocation half loops at the surface should be negligible at typical growth temperatures. On the other hand, the heterogeneous nucleation of half loops is much more likely. Here, heterogeneous nucleation refers to the nucleation of a half loop at an existing crystal defect, such

as a dislocation, void, precipitate, or scratch. The local strain field associated with such a defect could greatly reduce the activation energy for the creation of a dislocation half loop, thereby allowing this process to occur at an appreciable rate.

By a consideration of various experimental studies of relaxation in strained semiconductor structures, Dodson[41] showed that the observed behavior could be explained by dislocation sources that (1) become active at an excess stress close to zero, (2) become active at a physical stress as small as 0.01% of the shear modulus, and (3) result in a relaxation rate that is proportional to the square of the excess stress. On the basis of these observations, it was argued that the initial stages of lattice relaxation are dominated by the heterogeneous nucleation of dislocations near impurity clusters or other sources of localized stress concentration.

Gao et al.[42] developed atomistic models for the heterogeneous nucleation of dislocations at surface ledges in the case of $Si_{1-x}Ge_x$/Si(001), and discussed the atomistic details related to the formation of 60° glide dislocations, V-shaped defects, and wedge defects.

Direct evidence of dislocation half-loop nucleation (e.g., in the form of transmission electron microscopy [TEM] micrographs) is lacking in the literature. This does not necessarily mean that this process is inactive. Instead, it may be an indication that super-critical half loops expand rapidly once nucleated. However, Zou et al.[43] presented experimental evidence from SiGe islands on Si(001) that suggests a mechanism of strain relaxation by the nucleation of partial dislocation half loops at the surface.

5.3.3 Dislocation Multiplication

In many heteroepitaxial semiconductors, the observed lattice relaxation can only be explained by invoking either the nucleation of new dislocations or dislocation multiplication. This has been shown by Beanland,[44] based on the work of Matthews et al.[45] Consider a heteroepitaxial layer on a square substrate with sides L, parallel to the misfit dislocation lines. If the threading dislocation density in the substrate is D, then the total number of dislocation sources is L^2D. If, in the process of lattice relaxation, misfit dislocations are produced along the two possible directions with equal numbers, then there will be $L^2D/2$ misfit dislocations in each direction. If these misfit dislocations run to the edge of the sample (an optimistic assumption), and their sources (the threading dislocations) are uniformly distributed across the sample, their average length will be $L/2$. The linear density of misfit dislocations in the interface will be

$$\rho = \frac{(L^2D/2)(L/2)}{L^2} = \frac{LD}{4}. \tag{5.64}$$

The amount of strain that can be relaxed by this density of misfit dislocations is

$$\delta = \rho\, b \sin\alpha \sin\phi = (LD/4)\, b \sin\alpha \sin\phi, \tag{5.65}$$

where:

- b is the length of the Burgers vector
- α is the angle between the Burgers vector and line vector
- ϕ is the angle between the interface and the normal to the slip plane

Therefore, the amount of strain that can be relieved by a uniform density of sources without multiplication is proportional to the linear size of the sample. This conclusion holds

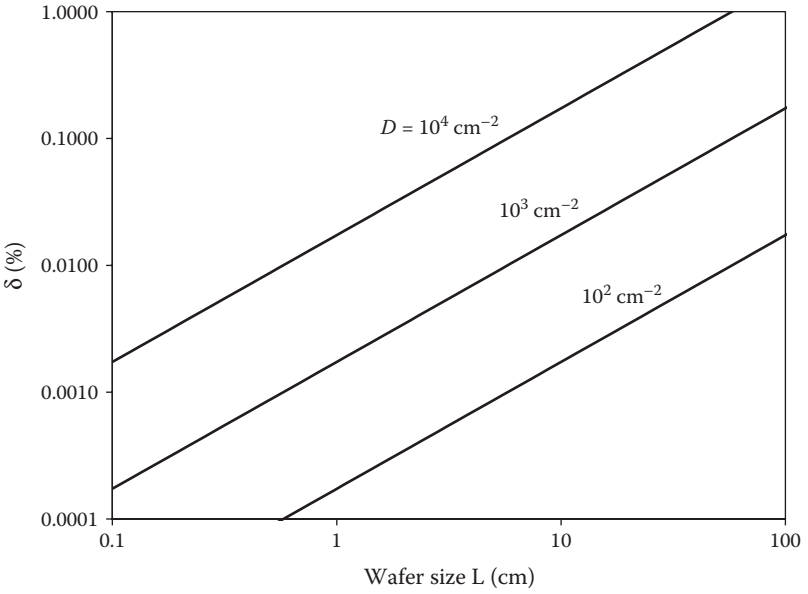


FIGURE 5.15
Maximum misfit that can be relieved by the bending over of existing threading dislocations, without dislocation multiplication, as a function of the substrate linear size. The substrate threading dislocation was assumed to be 10^2, 10^3, and 10^4 cm^{-2}, as indicated.

true for any substrate shape, although geometrical factors must be included in the analysis. For example, in the (001) heteroepitaxy of zinc blende semiconductors, a maximum of 1% mismatch strain may be relieved with a substrate dislocation density of 10^5 cm^{-2}, in the absence of dislocation multiplication. Figure 5.15 shows the amount of misfit strain that may be relieved as a function of the substrate size, with the substrate threading dislocation as a parameter, for the (001) heteroepitaxy of a zinc blende semiconductor with 60° dislocations on {111} glide planes. A square wafer is assumed, and the values of threading dislocation density considered are 10^2, 10^4, and 10^5 cm^{-2}, which encompass the typical range for practical substrates for heteroepitaxy. This figure shows that, in order to account for the observed lattice relaxation in heteroepitaxial semiconductors, we must invoke either dislocation multiplication or an extremely high density of sources for the heterogeneous nucleation of dislocations. Dislocation nucleation sources other than substrate threading dislocations, if present, are unlikely to have a density much greater than D in carefully prepared, high-quality substrates. Either a single such source must produce many dislocations, or dislocation multiplication must be important in the lattice relaxation of heteroepitaxial semiconductors.

5.3.3.1 Frank–Read Source

One possible mechanism for the multiplication of dislocations is the Frank–Read source,[46] illustrated in Figure 5.16. Here, the preexisting dislocation is anchored at points D and D′. It is important to note that the dislocation may not end within a perfect crystal, so the points D and D′ represent bends in the crystal dislocation, rather than actual terminations. There are many possible reasons why the dislocation could be immobile at points D and D′. Whereas the plane of the paper is assumed to be a glide plane, the bent-over portions

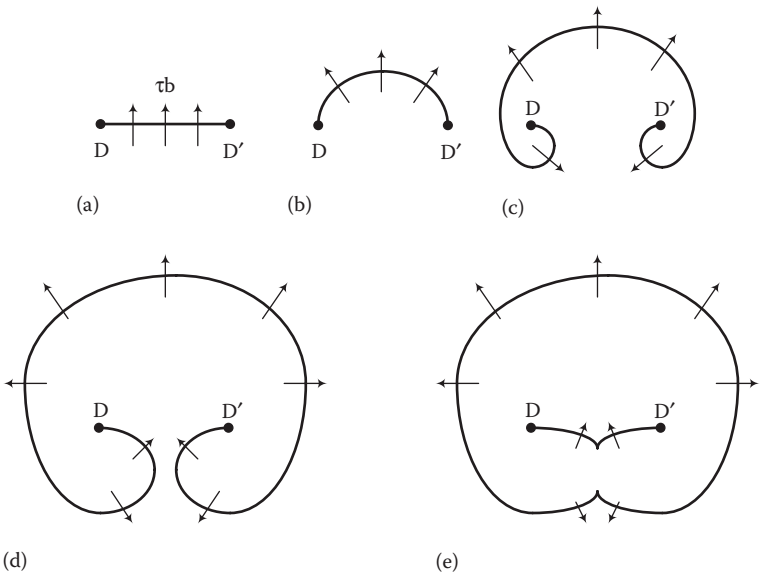


FIGURE 5.16
Frank–Read source.

of the dislocation may not lie on easy glide planes, rendering them essentially sessile. Another possibility is that defects could pin the dislocation at points D and D′. These pinning defects could be inclusions, voids, or even other dislocations, but the important feature is that they immobilize the dislocation at these two points.

Regardless of these details, an applied stress, as shown in Figure 5.16a, will cause the dislocation to bow, as shown in Figure 5.16b. Eventually, the dislocation can start to bend back upon itself, as shown in Figure 5.16c. It should be recognized that whereas the Burgers vector is conserved along the length of the dislocation, the line vector is reversed on the trailing edge of the bowed dislocation relative to the leading edge. Therefore, the leading and trailing edges experience forces of opposite sign as shown, tending to further expand the bowing dislocation, as in Figure 5.16d. Eventually, the bowing dislocation closes upon itself as in Figure 5.16e. The dislocation loop so created can continue to expand under the applied stress. The dislocation segment between the pinning defects can now snap back into its original configuration, and the multiplication process can repeat. Hence, such a Frank–Read source can continue to eject dislocation loops as long as the necessary stress is applied.

Frank–Read dislocation sources have been observed experimentally by Dash and Meieran. Dash decorated dislocations in a Si crystal using a Cu precipitation technique, which rendered them visible by infrared transmission microscopy.[47] The roughly concentric hexagonal dislocation loops observed in the sample were attributed to a Frank–Read-type source. Meieran observed a Frank–Read source in a Si crystal using x-ray topography (Figure 5.17).

Beanland[44] considered the operation of Frank–Read sources in mismatched heteroepitaxial layers. A possible configuration for a Frank–Read source in this situation is shown in Figure 5.18. In Figure 5.18a, the threading dislocation is anchored at points A and B. With an applied stress, the dislocation bows out between A and B, as shown in Figure 5.18b. Upon reaching the surface, the bowing loop breaks into two dislocations, as in Figure 5.18c.

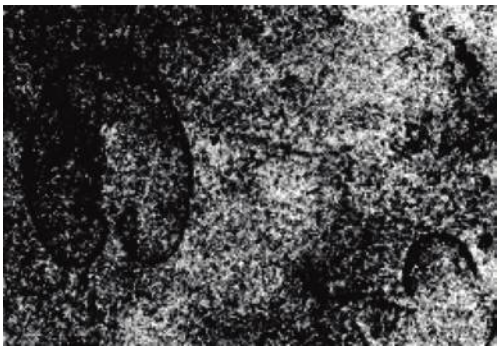


FIGURE 5.17
X-ray topograph ($[2\bar{2}0]$ reflection) of a sawed and chemically polished Si wafer, showing a bowing Frank–Read dislocation source. The magnification is ×15. (Reprinted from E. S. Meieran, *J. Appl. Phys.*, 36, 1497 [1965]. With permission. Copyright 1965, American Institute of Physics.)

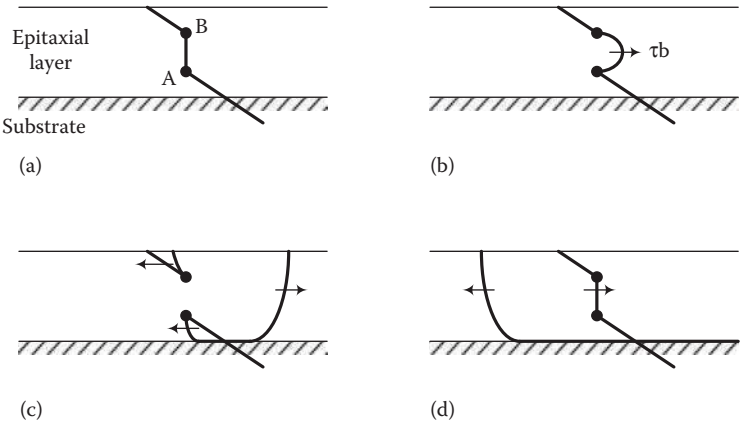


FIGURE 5.18
Possible configuration for a Frank–Read source in a heteroepitaxial layer. (a) The threading dislocation is anchored at points A and B. (b) With an applied stress, the dislocation bows out between A and B. (c) Upon reaching the surface, the bowing loop breaks into two dislocations. (d) Interaction of the two dislocations results in the formation of a half loop (the right side of which has glided out of the picture), leaving a dislocation similar to the original source defect. (Adapted from R. Beanland, *J. Appl. Phys.*, 72, 4031 [1992]. With permission. Copyright 1992, American Institute of Physics.)

Finally, interaction of the two dislocations results in the formation of a half loop (the right side of which has glided out of the picture), as shown in Figure 5.18d. This process leaves the original dislocation intact, and it can participate in further multiplication.

The critical thickness for the operation of such a Frank–Read source has been calculated by Beanland.[44] In this treatment, a force balance relationship was applied for the bowing of the pinned dislocation. The critical thickness so determined depends on the positions of the pinning points and the orientation of the pinned segment. However, assuming the pinned segment AB lies along the [112] direction, the minimum thickness for which the Frank–Read source may operate h_f is given by

$$h_f = h_c + 2h_p\,, \tag{5.66}$$

where:

h_c is the Matthews and Blakeslee critical layer thickness

h_p is the critical height for the formation of a spiral source given by

$$h_p = \frac{(2+\nu)b}{4\pi|f|(1-\nu)}\left[\ln\left(\frac{4\sqrt{6}h_p}{b}\right)+\frac{(\nu-2)}{(\nu+2)}\right], \tag{5.67}$$

where:

$f=(a_s-a_e)/a_e$ is the lattice mismatch strain

ν is the Poisson ratio

b is the length of the Burgers vector

Figure 5.19 shows the critical thickness for Frank–Read multiplication as a function of the lattice mismatch strain. Also shown is the Matthews and Blakeslee critical layer thickness for lattice relaxation, and the critical thickness for multiplication by spiral sources (considered in Section 5.3.3.2). Typically, h_f is four to seven times the Matthews and Blakeslee critical layer thickness. In many cases, a large fraction of the observed relaxation occurs after the thickness is several times the critical layer thickness. It is therefore likely, based on Beanland's estimates, for Frank–Read multiplication to be active in mismatched heteroepitaxial layers.

Frank–Read-type sources have been observed in heteroepitaxial layers by a number of workers. Lefevbre et al.[48] observed such a source in InGaAs/GaAs(001). LeGoues et al.[49] reported the observation of Frank–Read-type sources in SiGe/Si(001). Capano et al.[50] observed regular cross-slip in SiGe/Si(001) by x-ray topography, and explained this as a result of the operation of Frank–Read-type sources.

5.3.3.2 Spiral Source

Another type of dislocation source is the spiral source, also proposed by Frank and Read.[46] Such a spiral source is shown in Figure 5.20. Here, the dislocation line ABC is bent out of

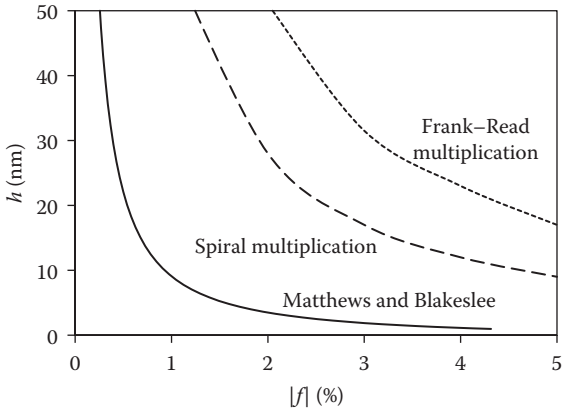


FIGURE 5.19
Critical thicknesses for dislocation multiplication by Frank–Read and spiral sources, as functions of the lattice mismatch strain $f=(a_s-a_e)/a_e$. Also shown is the Matthews and Blakeslee critical layer thickness for lattice relaxation, for comparison. (Adapted from R. Beanland, *J. Appl. Phys.*, 72, 4031 [1992]. With permission. Copyright 1992, American Institute of Physics.)

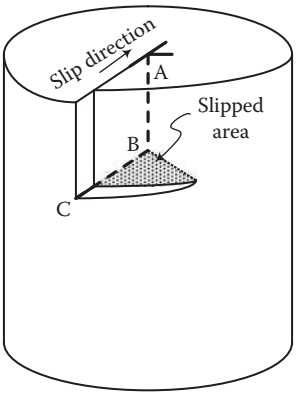


FIGURE 5.20
Bent dislocation line ABC, which can act as a spiral source. (Adapted from F. C. Frank and W. T. Read, *Phys. Rev.*, 79, 722 [1950]. With permission. Copyright 1950, American Physical Society.)

the horizontal glide plane at point B. Suppose the segment BC lies on a glide plane but segment AB is sessile. If a shear stress is applied on the glide plane and in the slip direction, the segment BC will sweep around the axis AB like the hand of a clock, producing one unit of slip for each revolution. Such a source is expected to sweep out a spiral, as shown in Figure 5.20, if glide is inhibited at the outer edge relative to the inner section. Although this mechanism does not produce new dislocations, it can increase the length of the dislocation line arbitrarily.

The spiral source has also been observed experimentally in Si by Dash[51] and Authier and Lang.[52] Figure 5.21 shows such a spiral source in a Si specimen. The image is an x-ray projection topograph obtained using the $\bar{1}\bar{1}\bar{1}$ reflection. The Si specimen was a rectangular bar, which was stressed by twisting about its long axis, which was [111], at 900°C.

A possible configuration for the spiral source in a heteroepitaxial layer was described by Beanland[44] and is shown in Figure 5.22. It is assumed that a threading dislocation is anchored at a single point A, as shown in Figure 5.22a. With an applied stress, the dislocation may bow out above the pinning point, as in Figure 5.22b. The bowed section will continue to expand and may glide to the interface to relieve mismatch strain, as in Figure 5.22c. Further expansion of the bowed portion may lead to production of a half loop if the bow reaches the surface and splits in two, as in Figure 5.22d. The original dislocation is then available to produce more dislocations by the same process. The main idea of spiral sources is that they use the substrate interface or the surface as pinning mechanisms.

The critical thickness for the operation of such a spiral source has been also calculated by Beanland.[44] This value depends on the position of the pinning point, but the minimum thickness for which the spiral source may operate h_s is given by

$$h_s = h_c + h_p, \tag{5.68}$$

where:

h_c is the Matthews and Blakeslee critical layer thickness
h_p is the height of the pinning point above the interface

In Figure 5.19, the critical thicknesses for multiplication by spiral and Frank–Read sources are compared. The spiral source can become active at two to four times the

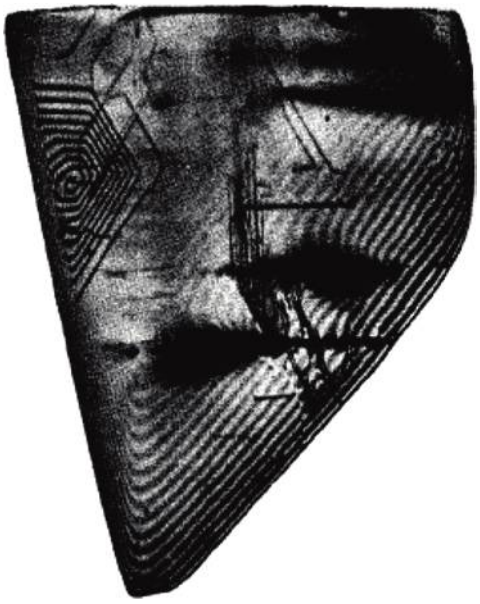


FIGURE 5.21
Spiral dislocation source in a Si crystal ($\bar{1}\bar{1}\bar{1}$ x-ray projection topograph). The approximate size of the spiral is 1.6 mm. (Reprinted from A. Authier and A. R. Lang, *J. Appl. Phys.*, 35, 1956 [1964]. With permission. Copyright 1964, American Institute of Physics.)

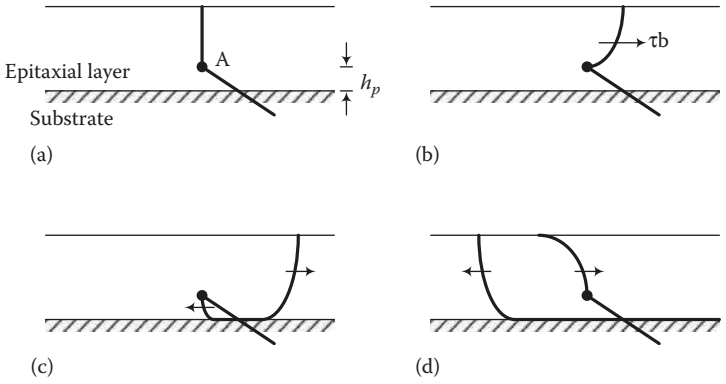


FIGURE 5.22
Possible configuration for a spiral source in a heteroepitaxial layer. (a) The threading dislocation is anchored at point A, which is located a distance h_p from the interface. (b) With an applied stress, the dislocation bows out above A. (c) The bowed section will continue to expand and may glide to the interface to relieve mismatch strain. (d) Upon reaching the surface, the bowing loop breaks into two dislocations. This results in the formation of a half loop (the right side of which has glided out of the picture), leaving a dislocation similar to the original source defect. (Adapted from R. Beanland, *J. Appl. Phys.*, 72, 4031 [1992]. With permission. Copyright 1992, American Institute of Physics.)

Matthews and Blakeslee critical thickness for lattice relaxation. It is therefore likely that both the spiral and Frank–Read mechanisms are active in relaxing heteroepitaxial layers.

Spiral sources have been seen by Mader and Blakeslee[53] in GaAsP/GaAs(113) and by Washburn and Kvam[54] in GeSi/Si(001).

5.3.3.3 Hagen–Strunk Multiplication

The previously considered mechanisms for multiplication involve the pinning of a dislocation at one (spiral source) or two (Frank–Read source) points in the presence of an applied stress. However, other multiplication mechanisms are possible that do not result from pinning of a dislocation but instead involve the intersection of two gliding dislocations. One such mechanism has been proposed by Hagen and Strunk[55] and is illustrated in Figure 5.23. Here, it is assumed that two dislocations AB and CD have the same Burgers vector but are on different glide planes. If these dislocations react at the cross-point, they may create two angular dislocations, as shown in Figure 5.23b. The repulsion of these dislocations with like Burgers vectors, in conjunction with the image force in a thin layer, can push the bent tip of one of the dislocations toward the surface on its inclined glide plane (e.g., a {111} plane for (001) zinc blende heteroepitaxy). Upon reaching the surface, the dislocation can split into two, as shown in Figure 5.23c. After a process involving the combined cross-slip and glide of the broken dislocation segments (Figure 5.23d), there will be three misfit dislocations, AE, FD, and COB, all of which may participate in further multiplication by the same process. The three misfit dislocations will all have the same Burgers vector and may appear in a configuration like that of Figure 5.23e.

Hagen et al. observed dislocations with a configuration similar to that shown in Figure 5.23e in heteroepitaxial Ge/GaAs(001) by TEM.[53,56] They interpreted these results as evidence for the operation of the Hagen–Strunk mechanism in heteroepitaxial layers. However, simple dislocation reactions at the interface could result in similar

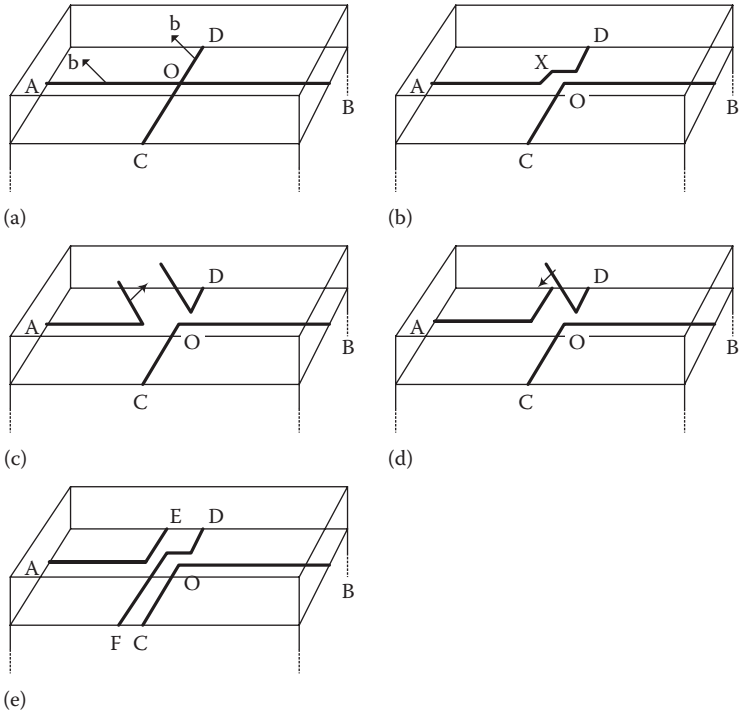


FIGURE 5.23
Hagen–Strunk dislocation multiplication mechanism. (Adapted from R. Beanland, *J. Appl. Phys.*, 72, 4031 [1992]. With permission. Copyright 1992, American Institute of Physics.)

configurations,[57] and so it has been argued that the TEM results do not provide evidence for Hagen–Strunk multiplication.[42]

Obayashi and Shintani[58] made a theoretical investigation of the Hagen–Strunk mechanism in Ge/GaAs(001) and SiGe/Si(001) systems. They considered the forces acting on a dislocation segment created by the reaction between two crossing misfit dislocations. They found that, because of the involvement of the image forces in this multiplication scheme, there is a critical thickness above which the mechanism cannot operate. They calculated the Hagen–Strunk critical thickness to be smaller than the Matthews and Blakeslee critical layer thickness for lattice relaxation. Based on this finding, they concluded that Hagen–Strunk multiplication is unlikely to occur in heteroepitaxial layers.

In summary, multiplication of dislocations in heteroepitaxial semiconductors is rather complex. Several mechanisms for dislocation multiplication have been proposed, including the Frank–Read source and spiral source. Frank–Read sources have been observed experimentally in both bulk and heteroepitaxial semiconductors. Similarly, there have been reports of observed spiral sources in both bulk semiconductors and mismatched layers. Theoretical calculations by Beanland point to the likelihood of these mechanisms being active in mismatched heteroepitaxy. Despite this, clear experimental observations of them remain extremely rare. Often, the dislocations in partially relaxed heteroepitaxial layers take on complex configurations that are difficult to interpret. One contributing factor might be that these multiplication sources are only able to operate a few times in highly dislocated crystals, whereas they might go unnoticed unless they operate many times. Still, the body of experimental and theoretical work indicates that sources of these types must be active during heteroepitaxial growth. That is, any configuration giving rise to the pinning of a dislocation at one or more points, in the presence of applied stress, should lead to dislocation multiplication.

It should be noted that other multiplication mechanisms are possible, which do not require the anchoring of existing dislocations, but instead involve the interaction of two or more dislocations. One such process that has been proposed is the Hagen–Strunk mechanism. There is very limited experimental evidence for this mechanism, and theoretical calculations show that this mechanism is unlikely in heteroepitaxial semiconductors. Still, there might be two-dislocation multiplication processes that are important but remain undiscovered. Further experimental investigations may lead to a better understanding of the complex processes involved in dislocation multiplication.

5.4 Interactions between Misfit Dislocations

In diamond and zinc blende (001) heteroepitaxial layers, the misfit dislocations usually have 60° or edge (90°) character and lie along the orthogonal ⟨110⟩ directions. There are three basic types of interactions that can occur at the intersections of these misfit dislocations,[59–61] which are shown in Figure 5.24. If the Burgers vectors are (a) parallel or (b) antiparallel, the intersecting dislocations will form two L-shaped dislocations. (c) If the Burgers vectors make an angle of 60°, a linking $a/2\langle 110\rangle$ dislocation will form. (d) If the Burgers vectors are perpendicular, no reaction is expected.

If only 60° misfit dislocations are present at the interface, for (001) diamond or zinc blende heteroepitaxy, then there are four possible Burgers vectors. This results in 16 possible interactions, 4/16 (25%) of which should produce L-shaped dislocations, 8/16 (50%) of which should produce linked dislocations, and 4/16 (25%) of which should result in no reaction.

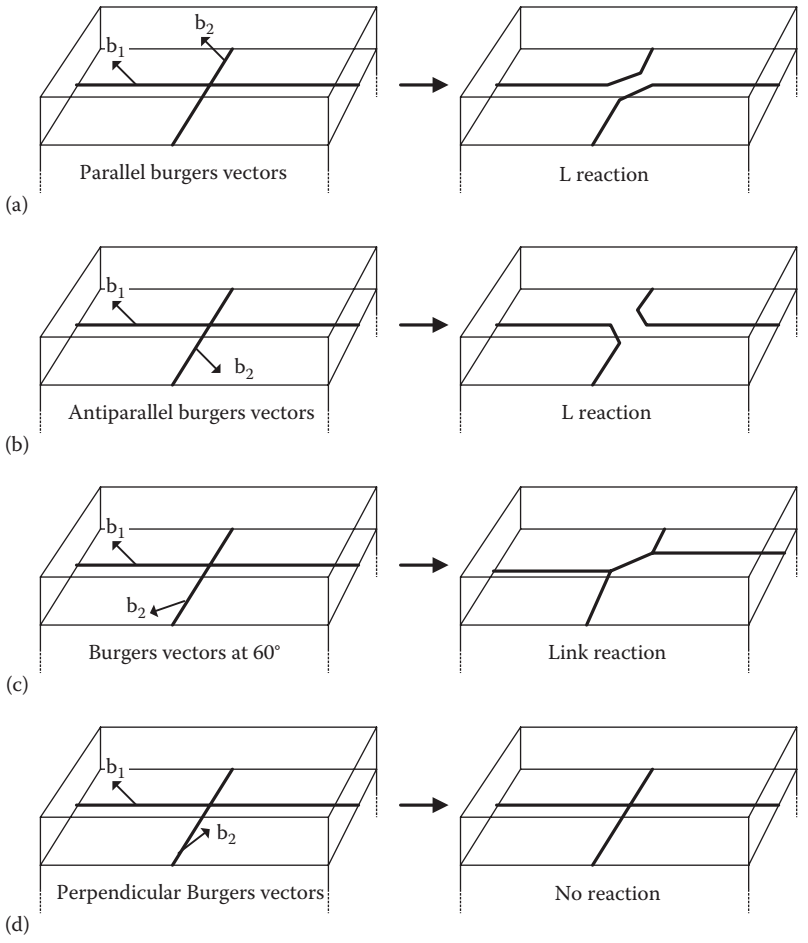


FIGURE 5.24
Three possible interactions between misfit dislocations along ⟨110⟩ directions in a diamond or zinc blende (001) heteroepitaxial layer. If the Burgers vectors are parallel (a) or antiparallel (b), the intersecting dislocations will form two L-shaped dislocations. (c) If the Burgers vectors make an angle of 60°, a linking $a/2\langle 110\rangle$ dislocation will form. (d) If the Burgers vectors are perpendicular, no reaction is expected.

If edge dislocations are present as well as 60° dislocations, then there are five possible Burgers vectors. This results in 25 possible interactions, 4/25 (16%) of which should produce L-shaped dislocations, 16/25 (64%) of which should produce linked dislocations, and 5/25 (20%) of which should result in no reaction. And, if the possible Burgers vectors are present in equal number in each type of dislocation, the observed dislocation interactions should closely follow these percentages above.

In TEM observations of interacting misfit dislocations, it may be difficult to distinguish between linked and unreacted dislocations. In the former case, the links are expected to be very short. This is because the slip planes of the linked dislocations lie out of the plane of the interface, and so they are unlikely to move apart significantly. Still, approximately 16% of the interactions should produce L-shaped dislocations if edge dislocations are excluded, but the inclusion of edge dislocations should result in L-shaped dislocations at a greater fraction of the intersections.

Dixon and Goodhew[57] examined about 1000 intersections between misfit dislocations in 20 nm thick $In_{0.2}Ga_{0.8}As/GaAs(001)$, grown by molecular beam epitaxy (MBE). They found L-shaped dislocations were present at 18% of the intersections, and interpreted this result as indicating the presence of some edge dislocations.

5.5 Lattice Relaxation Mechanisms

Some simple lattice relaxation mechanisms have already been considered briefly, in the sections on the critical layer thickness, dislocation nucleation, and dislocation multiplication. For example, Matthews and Blakeslee considered the bending over of substrate dislocations in deriving the critical layer thickness. However, as has been shown in Section 5.3.3, this mechanism alone cannot account for the measured extent of lattice relaxation in most heteroepitaxial systems. This leads us to invoke mechanisms involving the nucleation of new dislocations, or dislocation multiplication, as considered in the previous sections. Further, in the growth of heteroepitaxial islands (Volmer–Weber growth mode) new relaxation mechanisms are possible, such as the injection of misfit dislocations at the island boundaries. The purpose of this section is to describe the lattice relaxation processes in more detail and to describe the resulting defect structures that are to be expected.

5.5.1 Bending of Substrate Dislocations

Practical substrates for heteroepitaxy typically contain threading dislocations with a density of 10 to $10^5\,cm^{-2}$. These dislocations are replicated in the epitaxial layer and can glide to create misfit dislocations at the interface. This is the mechanism considered by Matthews and Blakeslee in their model for the critical layer thickness.

This lattice relaxation mechanism is shown schematically in Figure 5.25. In Figure 5.25a, a substrate threading dislocation AO has replicated in the epitaxial layer. If the layer is sufficiently thick ($h > h_c$), the threading segment will glide under the influence of the misfit stress, creating a misfit segment OC, as shown in Figure 5.25b. The dislocation will continue to glide, as shown in Figure 5.25c, increasing the total length of the misfit segment and reducing the average strain, unless it is impeded by a pinning defect or other dislocation. In rare circumstances, the threading segment may glide all the way to the wafer edge, annihilating the threading segment in the epitaxial layer, as shown in Figure 5.25d. Usually, however, the original threading segment BC will remain in the epitaxial layer. It is important to note that *if only this mechanism is active, the epitaxial layer must have a threading dislocation density equal to or less than that of the starting substrate.*

The amount of lattice mismatch that may be relieved by this mechanism depends on the substrate dislocation density and the average length for the misfit segments of the dislocations.

First, suppose that the substrate is square with sides of length L, parallel to the misfit dislocation lines. If the threading dislocation density in the substrate is D, then the total number of dislocations to be bent over is L^2D. If, in the process of lattice relaxation, misfit dislocations are produced along the two possible directions with equal numbers, then there will be $L^2D/2$ misfit dislocations in each direction. If these misfit dislocations run to the edge of the sample (an optimistic assumption), and their sources (the threading

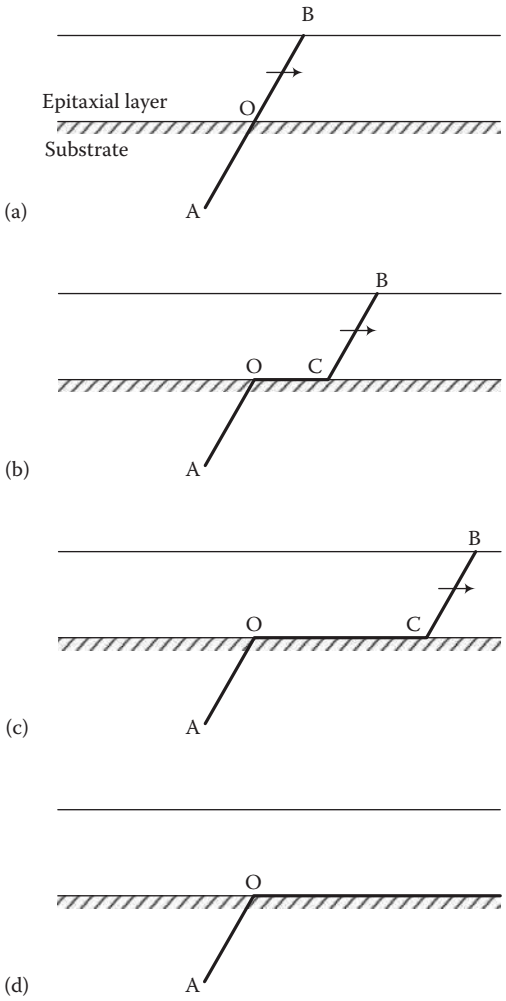


FIGURE 5.25
Relaxation mechanism involving the bending over of an existing substrate dislocation. (a) The substrate threading dislocation AO is replicated in the epitaxial layer. (b) If the layer thickness exceeds h_c, the threading segment will glide under the influence of the misfit stress, creating a misfit segment OC. (c) Unless impeded by a pinning defect or other dislocation, the threading segment will continue to glide to the right, increasing the length of the misfit dislocation and relaxing the mismatch strain in the epitaxial layer. (d) In rare circumstances, the threading segment may glide all the way to the wafer edge, annihilating the threading segment in the epitaxial layer. Otherwise, the threading segment BC will remain in the epitaxial layer.

dislocations) are uniformly distributed across the sample, their average length will be $L/2$. The linear density of misfit dislocations in the interface will be

$$\rho = \frac{\left(L^2 D/2\right)(L/2)}{L^2} = \frac{LD}{4}. \tag{5.69}$$

The amount of strain that can be relaxed by this density of misfit dislocations is

$$\delta = \rho\, b \sin\alpha \sin\phi = (LD/4)\, b \sin\alpha \sin\phi\,, \tag{5.70}$$

where:

- b is the length of the Burgers vector
- α is the angle between the Burgers vector and line vector
- ϕ is the angle between the interface and the normal to the slip plane

Typically, impediments to dislocation glide will limit the lengths of the misfit segments to be much less than the size of the substrate. If the average length for a misfit segment is L_{ave}, then the amount of lattice mismatch strain that can be relieved by bending over all of the threading dislocations is

$$\delta = (DL_{ave}/2)\, b \sin\alpha \sin\phi\,. \tag{5.71}$$

For the (001) heteroepitaxy of a zinc blende semiconductor,

$$\delta \approx (1\times10^{-8}\,\text{cm})DL_{ave}\,. \tag{5.72}$$

So, with an average misfit segment length of 100 μm and a substrate threading dislocation density of $10^5\,\text{cm}^{-2}$, only 0.001% mismatch strain may be relieved by this mechanism.

5.5.2 Glide of Half Loops

In most cases, the observed extent of lattice relaxation cannot be explained solely on the basis of bending of substrate dislocations, even though this may be the first mechanism to become active. Instead, it is necessary to invoke mechanisms involving the nucleation of new dislocations or dislocation multiplication. Even though homogeneous nucleation of dislocation half loops is not expected, their heterogeneous nucleation at a surface defect or concentration of stress is likely to occur, and dislocation half loops have been observed in the case of GaN grown on 6H-SiC (0001) by metalorganic vapor phase epitaxy (MOVPE) with an intermediate AlN buffer.[62] Such a half loop can expand to create a length of misfit dislocation, as illustrated in Figure 5.26. Suppose that a dislocation half loop ABCD is nucleated at a defect or region of concentrated stress at the surface, as shown in Figure 5.26a. This half loop can expand by the glide of its segments AB, BC, and CD, as shown in Figure 5.26b. By continued expansion, the half loop may reach the interface, as shown in Figure 5.26c, resulting in a misfit segment BC, as well as two threading segments AB and CD. Lattice relaxation can continue by the expansion of the half loop, as shown in Figure 5.26d. In rare circumstances, one of the threading segments may glide all the way to the wafer edge and annihilate. In this case, only one threading segment will remain in the epitaxial layer; otherwise, there will be two threading segments associated with each misfit segment. In contrast to the bending over of substrate dislocations, *the half-loop mechanism causes an increase in the threading dislocation density compared with that in the starting substrate.*

If a mismatched heteroepitaxial layer is completely relaxed by the glide of half loops, and the average size of the half loops (i.e., the average length of their misfit segments) is L_{ave}, and each misfit segment has two threading segments associated with it, then the threading dislocation density will be

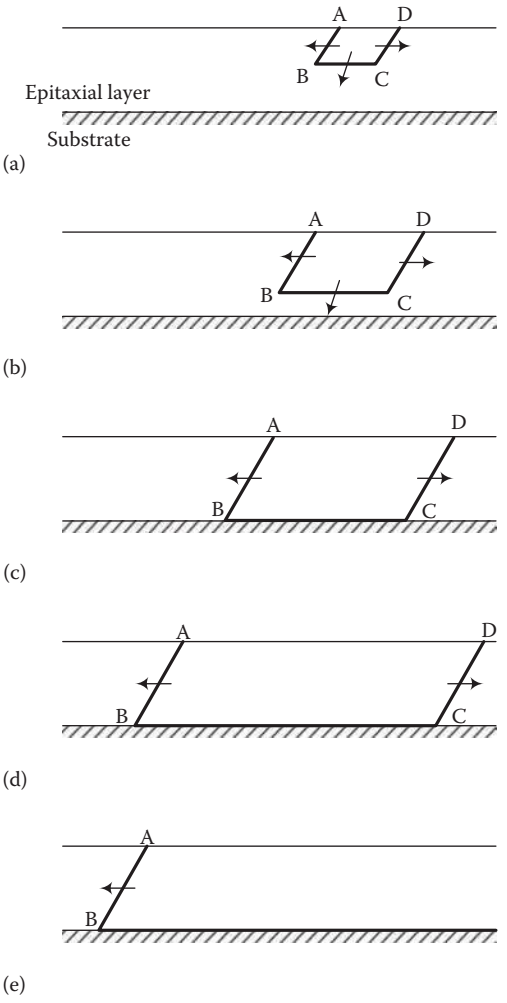


FIGURE 5.26
Relaxation mechanism involving the nucleation and glide of a half loop. (a) A dislocation half loop ABCD is nucleated at a defect or region of concentrated stress at the surface. (b) In response to the mismatch stress, the loop can expand on its glide plane, by the glide of the segments AB, BC, and CD. (c) The half loop may reach the interface by continued expansion, resulting in a misfit segment BC, as well as two threading segments AB and CD. (d) The half loop can continue to expand, lengthening its misfit segment and relaxing mismatch strain in the process. (e) In rare circumstances, one of the threading segments may glide all the way to the wafer edge and annihilate. In this case, only one threading segment will remain in the epitaxial layer; otherwise, there will be two threading segments associated with each misfit segment.

$$D = \frac{|f|}{L_{\text{ave}} b \sin\alpha \sin\phi}, \tag{5.73}$$

where b is the length of the Burgers vector, α is the angle between the Burgers vector and line vector, and ϕ is the angle between the interface and the normal to the slip plane. Assuming 60° misfit dislocation segments in a (001) zinc blende semiconductor, with an

average half loop width of 100 μm, the relaxation of 1% lattice mismatch will result in a threading dislocation density of about $5 \times 10^7\,cm^{-2}$.

5.5.3 Injection of Edge Dislocations at Island Boundaries

Many highly mismatched heteroepitaxial layers grow in a Volmer–Weber (three-dimensional) mode. In such a case, pure edge misfit dislocations can be injected at the boundaries of the growing islands, prior to island coalescence, and then glide on the interfacial plane. This phenomenon has been observed in a number of highly mismatched zinc blende (001) heterointerfaces by high-resolution TEM. The presence of these edge dislocations cannot be explained by nucleation and glide from the sample surface, because the slip direction for these dislocations lacks a component perpendicular to the interface. Another possible mechanism is the reaction of two 60° dislocations at the intersection of their glide planes. This Lomer–Cottrell mechanism is likely to be active, but it cannot account for the high numbers of edge dislocations observed in some heteroepitaxial systems. Climb is another possible explanation for the introduction of edge dislocations, but requires long-range diffusion and is expected to proceed too slowly at typical growth temperatures to explain the experimental observations.

It might be expected that layers relaxing by this mechanism would exhibit low threading dislocation densities. If the strain is relaxed by edge dislocations in this manner, there is no need for the nucleation of dislocations at the surface of the growing layer. At the same time, it is not possible for the edge misfit dislocations to glide upward toward the film surface. However, if the misfit strain is not fully relaxed at the time of island coalescence, then further relaxation may proceed by the glide of dislocation half loops from the surface, accompanied by the introduction of threading dislocations. Additionally, the matching of atomic bonds at the region of coalescence between neighboring islands may introduce "geometrically necessary" dislocations during the process of coalescence, and these can thread to the film surface. For these reasons, heteroepitaxial layers growing by a Volmer–Weber mechanism typically have large threading dislocations.

A material system that has been found to exhibit this mechanism of lattice relaxation under some conditions is GaSb/GaAs(001), which was studied by Qian et al.[63] They examined the interfacial misfit (IMF) dislocations in MBE-grown structures using high-resolution TEM. They found an array of pure edge dislocations with $b = \pm a/2[110]$ having a spacing of 57 ± 2 Å along each ⟨110⟩ direction. Within the experimental error, this is equal to the spacing of 55 Å at which the edge dislocations would completely relieve all of the mismatch strain ($f = -7.78\%$). Figure 5.27 shows a high-resolution TEM lattice image of the GaSb/GaAs(001) interface along the [110] direction. Each edge dislocation is associated with two extra {111} half planes in the GaAs substrate as marked.

In a separate experiment, Qian et al.[64] studied the initial stages of relaxation in GaSb/GaAs(001) grown by MBE. They found the edge dislocations existed in the growing islands, prior to coalescence. The misfit dislocations in the interior part of each island had a uniform spacing, but the spacing of the outermost dislocation was typically larger. The suggested interpretation of this observation was that the misfit dislocations nucleate at the leading edges of the {111} planes of the islands and then glide inward on the (001) plane; that is, the 90° misfit dislocations are injected at the advancing boundaries of the islands.

Another mechanism of relaxation by edge dislocations in GaSb/GaAs has also been observed in continuous thin films, and is referred to as an IMF dislocation mode of growth; this is discussed in Section 5.5.6.

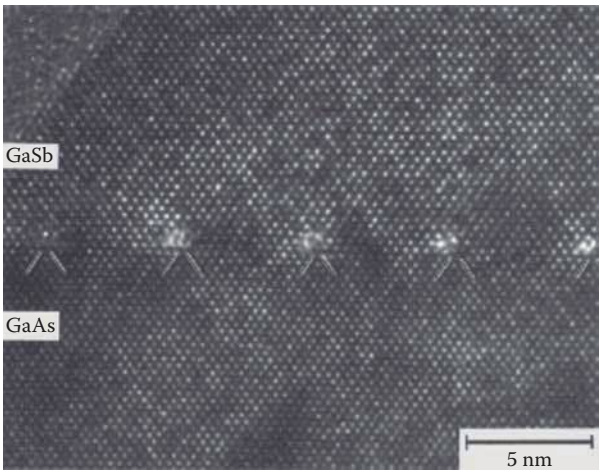


FIGURE 5.27
High-resolution TEM lattice image of the GaSb/GaAs(001) interface along the [110] direction. Each edge dislocation is associated with two extra {111} half planes in the GaAs substrate as marked. (Reprinted from W. Qian et al., *J. Electrochem. Soc.*, 144, 1430 [1997]. With permission. Copyright 1997, Electrochemical Society.)

5.5.4 Nucleation of Shockley Partial Dislocations

Petruzzello and Leys[14] found differences in the misfit dislocation structure between tensile and compressive interfaces in GaP/GaAsP and GaAsP/GaP interfaces, which they attributed to lattice relaxation by the nucleation of partial dislocations. Moreover, they showed theoretically that this mechanism results in differences in the lattice relaxation behavior in tensile versus compressive layers.

In their work, Petruzzello and Leys investigated the misfit dislocation structure at interfaces having both signs of mismatch strain in a GaP/$GaAs_{0.3}P_{0.7}$/GaP 001 heterostructure grown by MOVPE. In this structure, the $GaAs_{0.3}P_{0.7}$ layer was 2600 Å thick and the GaP cap was 900 Å thick. The room temperature lattice mismatch strains at these interfaces were ±1.1%, corresponding to a Matthews and Blakeslee critical layer thickness of ~8 nm. (The mismatch strain is compressive for the $GaAs_{0.3}P_{0.7}$/GaP interface and tensile for the GaP/$GaAs_{0.3}P_{0.7}$ interface.)

At the tensile interface (positive mismatch strain), Petruzzello and Leys found a square grid network of perfect and partial dislocations aligned with the ⟨110⟩ directions. These observations are consistent with the nucleation of partial dislocations in the mismatched layer with tensile strain. At the compressive interface (negative mismatch strain), however, the network of misfit dislocations involved only perfect dislocations, some of which were curved. This might indicate the involvement of a cross-slip mechanism that can only occur with perfect dislocations.

Petruzzello and Leys explained these differences between the compressive and tensile layers using the model of Marée et al.[65] for relaxation by the nucleation of Shockley partial dislocations. In a zinc blende heteroepitaxial layer, a 30° partial and a 90° partial may nucleate and then react at the interface to produce a 60° misfit dislocation. For example, in a layer with (001) orientation, if the lines of the dislocations are parallel to the $[1\bar{1}0]$ direction, the Burgers vectors for the 90° and 30° partials could be $a/6[112]$ and $a/6[211]$, respectively. These Shockley partials can react to form a single 60° misfit dislocation by the reaction

$$\frac{a}{6}[112]+\frac{a}{6}[211]\rightarrow\frac{a}{2}[101]. \tag{5.74}$$

In a tensile layer, the 90° partial will nucleate first, followed by the 30° partial. In the compressive layer, the partials are nucleated in the reverse order. This can be shown by consideration of the atomic arrangement on the {111}-type planes, shown schematically in Figure 5.28. The solid circles represent atoms in a layer of the {111}-type plane, and the dashed circles represent atoms in the underlying layer. The Burgers vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}$ are for the 90° partial, 30° partial, and 60° perfect dislocation, respectively. $\mathbf{u}$ represents the line of the dislocations. The direction of the resolved shear stress τ corresponds to the tensile case. For the situation shown, the slip of atoms in the layer by $\mathbf{b}_1$ will bring them to low-energy positions over the voids in the underlying layer, but the same is not true for slip by the partial Burgers vector $\mathbf{b}_2$. Therefore, in the tensile case, the 90° partial will nucleate first. Following this, the 30° Shockley partial will nucleate, with a stacking fault existing between the two partials. The 30° partial will glide toward the 90° partial, and they will eventually react to annihilate the stacking fault and form a perfect 60° dislocation.

Following the same arguments, we expect the 30° partial to nucleate first in the layer with compressive stress, in which the sign of τ is reversed. However, negligible dissociation is expected in this case because of the greater force on the 90° partial (whose Burgers vector is parallel to τ) compared with that on the 30° partial (whose Burgers vector is at a 60° angle to τ). In the compressive case, therefore, we expect that the lattice relaxation will occur predominantly by the glide of perfect 60° dislocations. The cross-slip of such a perfect dislocation from one {111} plane to another as it glides to the interface can result in the curved dislocation lines that Petruzzello and Leys observed at the compressive interface.

It should be noted that this mechanism of relaxation by Shockley partial dislocations does not alter the Matthews and Blakeslee critical layer thickness for the bending over of threading dislocations from the substrate. However, it could affect the critical layer thickness for the nucleation or multiplication of dislocations, and therefore the observable critical layer thickness. Since the more stressed 90° partial is nucleated first in the tensile case, this difference might cause the measurable onset of relaxation to occur in tensile layers with a smaller thickness than in compressive layers.

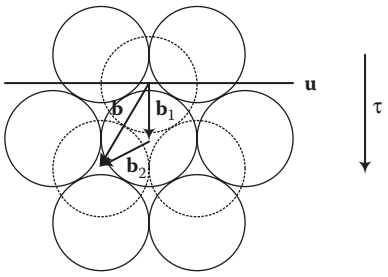


FIGURE 5.28
Atomic arrangement of the {111} planes of diamond and zinc blende semiconductors. The Burgers vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}$ are for the 90° partial, 30° partial, and 60° perfect dislocation, respectively. $\mathbf{u}$ represents the line of the dislocations. The resolved shear stress on the plane is τ, and the direction shown corresponds to the case of tensile stress. (Adapted from J. Petruzzello and M. R. Leys, *Appl. Phys. Lett.*, 53, 2414 [1988]. With permission. Copyright 1988, American Institute of Physics.)

5.5.5 Cracking

Another lattice relaxation mechanism is cracking, which has been observed in the case of wurtzite III-nitride semiconductors grown with tensile mismatch strain. Ito et al.[66] studied the lattice relaxation of $Al_xGa_{1-x}N/GaN(0001)$. For this heteroepitaxial system, the lattice mismatch strain is positive (tensile strain) and given by $f \approx x(3.5\%)$ at room temperature. Ito et al. found that the tensile AlGaN layers exhibited cracking if the critical layer thickness was exceeded. They showed that this cracking resulted from the lattice relaxation mechanism, rather than the thermal strain introduced during cooldown. Cracking cannot relieve mismatch strain in compressive films, however. Therefore, the lattice relaxation by cracking in the tensile layers is indicative of a fundamental difference in lattice relaxation mechanisms between the tensile and compressive cases.

5.5.6 Interfacial Misfit Dislocation Growth Mode

In some highly mismatched heteroepitaxial layers, the misfit dislocations form in the early stages of epitaxy and tend to have a very regular spacing at the interface. It has also been noted that there is a tendency for 90° dislocations in high-mismatch ($|f| > 6\%$) systems, whereas moderate-mismatch ($3\% < |f| < 4\%$) results in mixed 60° and 90° dislocations, while the relaxation of low-mismatch material systems ($|f| < 2\%$) is dominated by 60° dislocations.[67] In some high-mismatch cases, the regularity of the IMF dislocation array, its tendency toward edge character, and the relative absence of threading dislocations, is so striking that it has been referred to as the IMF dislocation growth mode. The most studied material system in this regard is GaSb on GaAs, for which the lattice mismatch is ~7.78%. Early work in this material system was motivated by the lack of high-quality, semi-insulating GaSb substrates, but epitaxial growth using MBE[68,69] resulted in high threading dislocation densities of 10^7 to $10^9\,cm^{-2}$ in layers up to 7 μm thickness. Despite these high threading dislocation densities, Ivanov et al.[69] showed by cross-sectional TEM that the interface between GaSb and GaAs contained a highly regular array of misfit dislocations with a spacing of 5.5 nm. These defects were found by TEM diffraction analysis to have Burgers vectors of the type $a/2[1\bar{1}0]$, indicating 90° edge character. Studies of three-dimensional, island growth of GaSb on GaAs(001) with an average thickness greater than 10 monolayers (ML) also revealed very regular arrays of edge or 60° misfit dislocations,[70–72] leading Thibado et al.[73] to conduct a study of the evolution of GaSb epitaxy on GaAs(001)–$c(4 \times 4)$. It is possible that the mechanism of relaxation was the injection of edge dislocations at the sides of islands in these studies.

Huang et al.[74] demonstrated for the first time the growth of relaxed GaSb on GaAs(001) with a regular array of misfit dislocations and low threading dislocation density. The misfit dislocations were found by high-resolution transmission electron microscopy (HRTEM) Burgers circuit analysis to have edge character with Burgers vectors of the type $a/2[110]$, and exhibited a spacing of 5.6 nm, as shown in Figure 5.29. This periodicity corresponds to 13 lattice spacings in GaSb and 14 lattice spacings in GaAs, as expected for complete relaxation of the 7.78% mismatch by edge dislocations. KOH etch pit density measurements revealed threading dislocation densities as low as $7 \times 10^5 cm^{-2}$, three to four orders of magnitude lower than typical GaSb/GaAs(001) heteroepitaxial layers. Zhou et al.[75] reported similar findings for MOVPE growth of GaSb on GaAs(001), and showed that IMF-grown GaSb on GaAs exhibited a high hole mobility of $553\,cm^2\,V^{-1}\,s^{-1}$ at room temperature with a hole concentration of $5.27 \times 10^6\,cm^{-3}$.

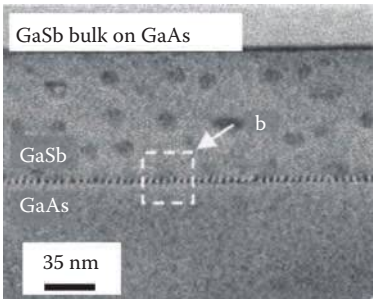


FIGURE 5.29
Cross-sectional TEM image of 120 nm GaSb/GaAs(001), showing a regular array of 90° misfit dislocations with a spacing of 5.6 nm. (Reprinted from S. H. Huang et al., *Appl. Phys. Lett.*, 88, 131911 [2006]. With permission. Copyright 2006, American Institute of Physics.)

In the case of GaSb/GaAs(001) heteroepitaxy, a proposed mechanism for the IMF growth mode involves the self-assembly of adsorbed Sb atoms on the GaAs(001) surface prior to the initiation of epitaxy.[71,76] It is postulated that exposure of the GaAs(001) surface to Sb for the appropriate time and temperature allows the Sb adatoms to form single as well as double bonds with the underlying Ga layer, without reacting with the underlying GaAs to displace As atoms. The existence of regular skipped bonds with underlying Ga atoms, occurring every 13th row of atoms along the [110] and $[1\bar{1}0]$ directions, is thought to give rise to misfit dislocations with edge character along these same two orthogonal directions. To the extent that the original self-assembly of adsorbed atoms formed a perfectly periodic structure, the misfit dislocations could extend from one end of the sample to the other, and the introduction of threading dislocations would be unnecessary.

The striking demonstration of IMF growth mode in compressive GaSb on GaAs(001) was followed by a similar demonstration of an IMF array in tensile GaAs on GaSb(001).[77] A difference in this case is the highly reactive nature of As_2. An arsenic soak of 60 s gives rise to etching of the GaSb surface, which results in nanovoids and a high threading dislocation density of $>10^9\,cm^{-2}$. A shorter arsenic soak of 10 s, though it eliminated most etch pits and nanovoids, still provided a high threading dislocation density of $10^7\,cm^{-2}$. On the other hand, good-quality GaAs with a threading dislocation density of $\sim 3 \times 10^6\,cm^{-2}$ was grown if no arsenic soak was employed, and a regular array of 90° misfit dislocations was seen by HRTEM, as shown in Figure 5.30.

Following this early work, there have been reports of an IMF growth mode in several material systems, including AlSb/Si(001),[78] InAs/GaAs(001),[79,80] and ternary and quaternary films on GaAs(001).[81] In a study of several material systems, Richardson et al.[81] noted that, on the basis of existing experimental evidence, there appear to be several pathways to IMF growth. For the IMF growth of GaSb on GaAs, Huang et al.[74] utilized an Sb soak, whereas Richardson et al. did not.[81] For IMF growth of InAs on GaAs, Eyink et al.[80] employed In-rich surfaces, whereas Richardson et al.[81] utilized As-rich conditions to initiate growth.

A proposed criterion for the evaluation of IMF growth is the ratio of the x-ray full widths at half maximum (FWHMs) from the reciprocal space map, $FWHM_x/FWHM_z$, where $FWHM_x$ is the width for the in-plane direction and $FWHM_z$ is the width for the out-of-plane (growth) direction.[78] Kaganer et al.[82] have calculated the theoretical reciprocal space maps for materials grown in an ideal IMF mode with a perfect array of edge dislocations at

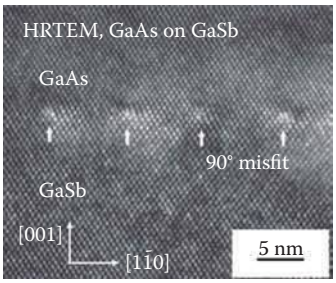


FIGURE 5.30
HRTEM image of IMF dislocation array in GaAs on GaSb(001). (Reprinted from S. H. Huang et al., *Appl. Phys. Lett.*, 90, 161902 [2007]. With permission. Copyright 2007, American Institute of Physics.)

the interface and absent of all threading dislocations. The reciprocal space map is expected to exhibit an elliptical peak with intensity proportional to

$$I(q_x, q_z) \propto \exp\left[-\frac{x_0(q_x - q_{x_0})^2}{4\sigma_x^2(0)} - \frac{(q_z - q_{z_0})^2}{4\sigma_z^2(0)}\right], \tag{5.75}$$

where $(q_x - q_{x_0})$ and $(q_z - q_{z_0})$ are the deviations from the reciprocal space point (q_{x_0}, q_{z_0}) in the in-plane and out-of-plane directions, respectively, and the standard deviations are given by

$$\sigma_x^2(0) = \frac{\rho Q_z^2 b_x^2}{8\pi h} \tag{5.76}$$

and

$$\sigma_z^2(0) = \left(\frac{\nu}{1-\nu}\right)^2 \frac{\rho Q_z^2 b_x^2}{4\pi h}, \tag{5.77}$$

where:

- ρ is the linear density of misfit dislocations
- ν is the Poisson ratio
- Q_z is the component of the reciprocal space vector in the growth direction
- b_x is the in-plane component of the Burgers vector for the edge dislocations
- h is the film thickness

The ratio of the widths in this ideal IMF case is therefore predicted to be

$$\frac{\text{FWHM}_x}{\text{FWHM}_z} = \frac{\sigma_x(0)}{\sigma_z(0)} = \frac{1-\nu}{\sqrt{2}\nu}, \tag{5.78}$$

which is equal to 1.574 for GaSb/GaAs(001), 1.313 for InAs/GaAs(001), and 1.257 for InP/GaAs(001). Whereas the absolute widths depend on the film thickness, the ratio of widths

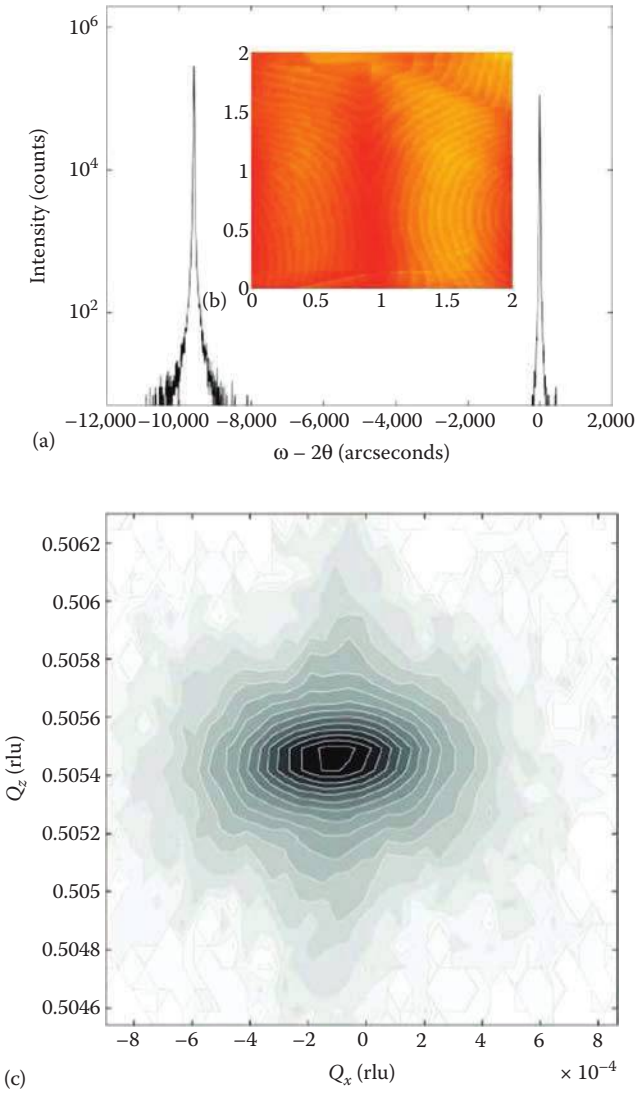


FIGURE 5.31
(a) 004 x-ray diffraction profile for a 5 μm thick layer of GaSb on GaAs(001), showing a FWHM of 22 arcsec. (b) AFM topograph for the same film, for which the horizontal scan direction is [110]. (c) Reciprocal space map around the 004 peak of the GaSb, showing the asymmetry in the in-plane (Qx) and out-of-plane (Qz) directions. (Reprinted from C. J. K. Richardson et al., *J. Vac. Sci. Technol. B*, 29, 03C126-1 [2011]. With permission. Copyright 2011, American Vacuum Society.)

does not. Figure 5.31 shows the 004 x-ray diffraction profile, AFM topograph, and x-ray reciprocal space map around the GaSb 004 reciprocal space point for a 5 μm thick GaSb layer on GaAs(001). This sample exhibits a width ratio of $FWHM_x/FWHM_z = 1.6$, which is equal to the theoretical value within the reported experimental error of ±10%. However, the absolute widths (22 arcsec in the growth direction) are greater than the theoretical values for a layer of this thickness, which may indicate that both $FWHM_x$ and $FWHM_z$ are broadened by threading dislocations.

In conclusion, IMF growth of highly mismatched materials has allowed a striking reduction in the threading dislocation density that should have important practical implications for device development. Threading dislocation densities in these materials have been demonstrated to be three to four orders of magnitude lower than for non-IMF growth in the same material system. On the other hand, the threading dislocation densities remain one to two orders of magnitude higher than the substrate materials, indicating that some threading dislocations are introduced along with the misfits. Additional investigation will be needed to understand and apply IMF growth to a wider range of materials.

5.6 Quantitative Models for Lattice Relaxation

Heteroepitaxial layers with low-mismatch strain ($|f| < 2\%$) will grow coherently strained ($\varepsilon_{\parallel}=f$) to match the lattice spacings of the substrate in the plane of the interface, up to the critical layer thickness h_c. Beyond the critical layer thickness, it becomes energetically favorable for the introduction of misfit dislocations to relieve some of the mismatch strain.

A number of models have been developed to describe the variation of the residual strain with film thickness in partially relaxed layers, which are greater than the critical layer thickness. Matthews and Blakeslee developed an equilibrium model that adequately describes the strain relaxation in heteroepitaxial layers for which there exist no significant kinetic barriers to the nucleation or glide of dislocations. Experimentally, however, it has been found that it is possible to grow metastable layers, with residual strains greatly exceeding those predicted by the equilibrium model. This has motivated the development of kinetic models for strain relaxation in mismatched heteroepitaxial layers.

The first such kinetic model appears to be that of Matthews et al., who modified the equilibrium theory with a term to account for the Peierls (lattice friction) force on moving dislocations. They made use of the dislocation velocity model developed by Haasen. However, it has been found that the Matthews, Mader, and Light model cannot accurately predict both the initial and later stages of the lattice relaxation, and this was thought to be because dislocation multiplication was not included in their model.

Dodson and Tsao developed a kinetic model that included a phenomenological model for dislocation multiplication, as well as an empirical model for dislocation glide under the influence of stress. This model has been used to successfully fit the relaxation characteristics of a number of heteroepitaxial layers from different material systems. Though the model involves two adjustable parameters, it appears to often provide a satisfactory description of the dislocation dynamics and lattice relaxation.

Hull and Bean[83] and Hull et al.[84] presented another model for kinetically limited strain relaxation, in which they accounted for dislocation nucleation but neglected dislocation multiplication. This model includes a stress-dependent activation energy for the glide of dislocations.[85] Therefore, the rate of relaxation depends superlinearly on the excess stress, as it does in the Dodson and Tsao model, although the physical origins of this dependence are considered to be different.

Kujofsa et al.[108,109] have subsequently generalized the Dodson and Tsao model to graded layers, multilayer structures, and layers grown under varying growth temperatures or subjected to growth interruptions or postgrowth annealing steps.

This section outlines the equilibrium and kinetic models described above. In each case, the starting assumptions and underlying physics are described, along with the resulting model equations. The practical applications of these models, and their limitations, are also summarized.

5.6.1 Matthews and Blakeslee Equilibrium Model

The Matthews and Blakeslee[2] equilibrium model is based on force balance for an existing threading dislocation, with the same physical basis as the Matthews and Blakeslee critical layer thickness. The resulting equilibrium strain in a heteroepitaxial layer of thickness h, with $h > h_c$, is given by

$$\varepsilon_{\parallel,\mathrm{eq}}(h) = \frac{f}{|f|}\frac{G\left(1-\nu\cos^2\alpha\right)}{8\pi h\left(1+\nu\right)\cos\lambda}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{5.79}$$

where:

- G is the shear modulus (assumed equal for the epitaxial layer and substrate)
- ν is the Poisson ratio
- α is the angle between the Burgers vector and the line vector for the dislocations
- λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface

The $f/|f|$ term takes on a value of ± 1 to account for the sign of the strain.

It is important to note that the equilibrium strain is (approximately) inversely proportional to the layer thickness. Therefore, heteroepitaxial layers of finite thickness will not relax completely even in equilibrium.

For application to the heteroepitaxy of zinc blende or diamond semiconductors with (001) orientation, we assume that the gliding dislocations are of the 60° type, with Burgers vectors of the type $a/2\langle 011\rangle$ and line vectors of the type $1/\sqrt{2}\,\langle 110\rangle$. The glide planes for these dislocations are {111}-type planes. Thus, $b = a/\sqrt{2}$, $\cos\alpha = 1/2$, and $\cos\lambda = 1/2$. Figure 5.32 shows the equilibrium strain versus the thickness for the heteroepitaxy of a diamond or zinc blende semiconductor.

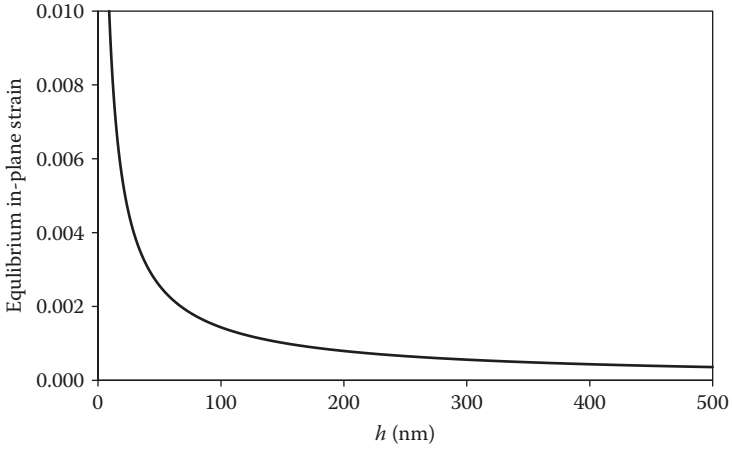


FIGURE 5.32
Equilibrium strain vs. thickness for a heteroepitaxial zinc blende layer with (001) orientation, calculated using the Matthews and Blakeslee model assuming $\cos\alpha = \cos\lambda = 1/2$, $b = 4.0$ Å, and $\nu = 1/3$.

5.6.2 Kinetic Models for Relaxation

Kinetic models consider the rates of physical processes, rather than just the final or minimum energy configuration. A key rate-limiting process is the glide of dislocations, and much work has been directed at developing an understanding of the flow of dislocations during mismatched heteroepitaxy. Generally, it is understood that the glide velocities are thermally activated and increase monotonically with excess stress.

Dislocation glide velocities have been measured by the "double-etch" method in GaAs,[86,87] Ge,[88,89] Si,[86] and InSb,[86] and by direct observation in Ge_xSi_{1-x}/Si using TEM.[90] In these materials, it has been found that the stress and temperature dependence of the dislocation glide velocity may be described by a phenomenological relationship,[91]

$$v = v_0 \exp\left(-\frac{\tau_0}{\tau}\right)\exp\left(-\frac{U}{k_B T}\right), \tag{5.80}$$

where:

- v_0 is a characteristic velocity
- τ_0 is a characteristic stress
- τ is the resolved shear stress on the glide plane in the slip direction
- U is the activation energy for glide
- T is the absolute temperature
- k_B is the Boltzmann constant

Over restricted ranges of stress, this relationship is commonly approximated by

$$v = B\tau^m \exp\left(-\frac{U}{k_B T}\right), \tag{5.81}$$

where m and B are constants.

Hirth and Lothe[92] have developed a physical model for the glide of dislocations based on the nucleation and propagation of kinks. Glide of the dislocation is accomplished through the propagation of many such single kinks.

According to this model, the nucleation rate for double kinks is

$$J = \frac{v_D b q \tau_{\text{eff}}}{k_B T}\exp\left(-\frac{2F_k}{k_B T}\right), \tag{5.82}$$

where:

- v_D is the attempt frequency, or Debye frequency (~$1.3 \times 10^{13}\ s^{-1}$ for Si)
- b is the length of the Burgers vector
- q is the distance between Peierls valleys (~0.33 nm in Si)
- τ_{eff} is the effective or excess stress (that part of the stress that is above that required to overcome the line tension in the misfit dislocation segment*)
- k is the Boltzmann constant

* In published work, the effective stress (or excess stress) has been denoted by four different symbols, which all mean the same thing: τ_{eff}, τ_{exc}, σ_{eff}, and σ_{exc}. In the descriptions, which follow, we will generally use the notation τ_{eff} for consistency.

T is the temperature
$2F_k$ is the energy required to nucleate a double kink ($F_k \approx 1$ eV in $Si_{1-x}Ge_x$)

The kink velocity of propagation is given by

$$v_k = \frac{v_D a^2 bq\tau_{\text{eff}}}{k_B T}\exp\left(-\frac{E_m}{k_B T}\right), \tag{5.83}$$

where:
a is the kink jump distance
E_m is the energy barrier for kink jumps along the line direction ($E_m \approx 1.2\,\text{eV} - x0.6\,\text{eV}$ for $Si_{1-x}Ge_x$)

Combining these relationships, the glide velocity for the dislocation is given by

$$v = \frac{2v_D abq^2\tau_{\text{eff}}}{k_B T}\left(\frac{L}{L+X}\right)\exp\left(-\frac{E_m + F_k'}{k_B T}\right), \tag{5.84}$$

where:
L is the length of propagating dislocations
X is the average separation between kinks, given by

$$X = 2a\exp\left(\frac{F_k'}{kT}\right); \tag{5.85}$$

and F_k' is the modified kink energy in the presence of stress, determined by consideration of the interaction between two kinks as described by Seeger and Schiller,[93] and given by

$$F_k' = F_k - \frac{1}{2}\sqrt{\frac{G(1+\nu)b^3q^3\tau_{\text{app}}}{2\pi(1-\nu)}}, \tag{5.86}$$

where τ_{app} is the applied (not the excess) stress. There are two limiting cases for the behavior.[94] In bulk crystals, $L \gg X$ and dislocations propagate by the motion of single kinks. The same is true for thin, uncapped epitaxial layers, in which single kinks may nucleate at the surface and propagate toward the interface as shown in Figure 5.33.

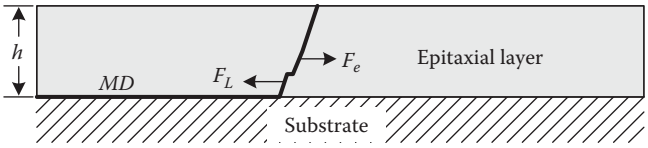


FIGURE 5.33
Glide of a threading dislocation by the single-kink mechanism in a thin, uncapped epitaxial layer.

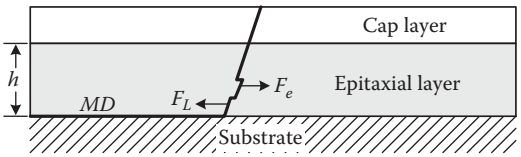


FIGURE 5.34
Glide of a threading dislocation by the double-kink mechanism in a capped epitaxial layer.

In the case of the single-kink mechanism, the glide velocity is given by

$$v = \frac{2v_D abq^2\tau_{\text{eff}}}{k_B T}\exp\left(-\frac{E_m + F_k'}{k_B T}\right) \quad (\text{single-kink model}). \tag{5.87}$$

In the other limit, applicable to capped epitaxial layers, dislocation glide is by the double-kink mechanism, shown in Figure 5.34, and the glide velocity is given by

$$v = \frac{v_D bq^2\tau_{\text{eff}}h}{k_B T}\exp\left(-\frac{E_m + 2F_k'}{k_B T}\right) \quad (\text{double-kink model}). \tag{5.88}$$

Thus, the diffusive kink model predicts a stress-dependent activation energy for the glide of dislocations and provides a physical description for its basis. In the case of $Si_{1-x}Ge_x$ alloys, the values of E_m and F_k are known ($E_m \approx 1.2\,\text{eV} - x0.6\,\text{eV}$, $F_k \approx 1.0\,\text{eV}$). For other heteroepitaxial semiconductors, it is usually the case that only a single activation energy is given for the glide of dislocations. In such cases, the glide velocity may be considered to vary superlinearly with the excess stress, as in the model of George and Rabier.[91] It should be understood that either type of model may be used to fit the experimental data over a range of excess stress, even though they are based on different physical descriptions of dislocation glide. At the present time, it is therefore not clear which physical description is more accurate.

5.6.3 Dislocation Blocking

The models for dislocation glide described above were developed with consideration of a single, isolated dislocation. However, when a gliding threading dislocation encounters another dislocation with an orthogonal misfit component, there is a hindering of the dislocation glide by a repulsive force that opposes the glide force due to the mismatch strain.[95] As a consequence, the gliding dislocation is forced to pass through a reduced thickness above the orthogonal dislocation. If this reduced thickness is less than the critical layer thickness associated with the remaining strain in the film, then motion of the gliding dislocation will be arrested, and the dislocation is said to be "blocked" or "pinned" by the misfit dislocation. Striking evidence of this phenomenon has been given in the case of a 80 nm GaAs/36.4 nm $In_{0.1}Ga_{0.9}As$/80 nm GaAs(001) heterostructure by Paine et al.[96] This structure was grown by MBE at 520°C and then annealed for 120 s at 900°C. The plan-view TEM micrograph of Figure 5.35 shows five threading dislocations that have been blocked by a single orthogonal dislocation. This same blocking mechanism has been shown to be important in other material systems, such as SiGe/Si(001).[80,81,97,98]

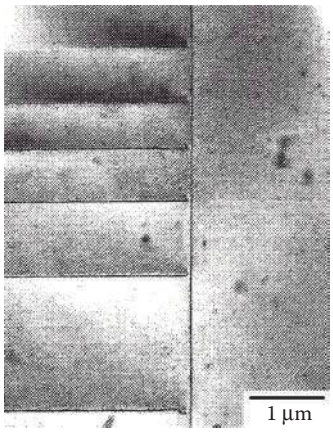


FIGURE 5.35
Plan-view TEM micrograph of a GaAs/$In_{0.1}Ga_{0.9}As$/GaAs(001) sandwich structure showing that five gliding dislocations have been blocked by a single orthogonal misfit dislocation. (Reprinted from L. B. Freund, *J. Appl. Phys.*, 68, 2073 [1990]. With permission. Copyright 1990, American Institute of Physics.)

Although the concept of dislocation blocking is easily understood, the quantitative analysis of the forces involved and the blocking criteria are quite complex. Freund[95] examined blocking in detail for heteroepitaxy of cubic materials having (001) orientation, and considering the various combinations of Burgers vectors for the gliding threading dislocation and the blocking misfit dislocation. This work was subsequently extended and compared with experimental results in the SiGe/Si(001) system by Gillard et al.[99] As shown in Figure 5.36, they considered a threading dislocation, with Burgers vector $\mathbf{b}_i$, gliding on the x_2, x_3' plane and encountering a misfit dislocation, with Burgers vector $\mathbf{a}_i$, lying parallel to

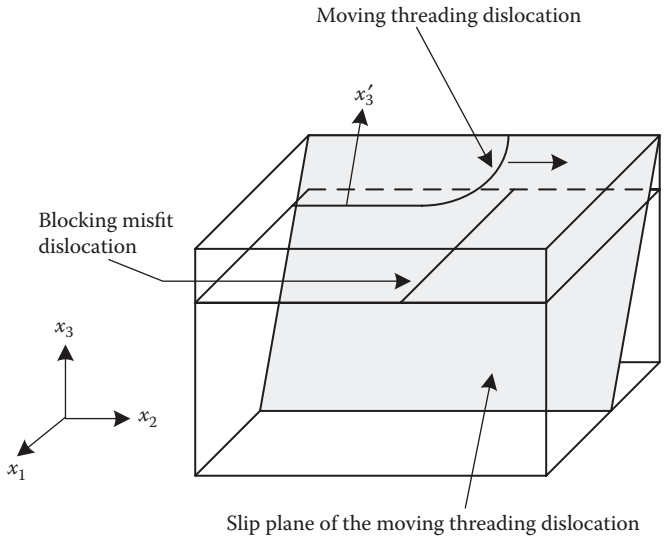


FIGURE 5.36
Gliding threading dislocation approaching an exisiting misfit dislocation in its path. (Adapted from V. T. Gillard et al., *J. Appl. Phys.*, 76, 7280 [1994]. With permission. Copyright 1994, American Institute of Physics.)

the x_1 axis. If the misfit dislocation is considered to have 60° character, then there are four independent combinations of Burgers vectors $\mathbf{a}_i/\mathbf{b}_i$.

The blocking behavior of the misfit dislocation may be described by the effect on the driving force for glide of the threading dislocation. This was considered in terms of the stress acting on the threading dislocation, $\sigma_{ij}^A + \sigma_{ij}^I$,
where:

σ_{ij}^A is the uniform applied stress associated with the residual strain
σ_{ij}^I is the stress associated with the misfit dislocation

The total glide force acting on the threading dislocation can be found as

$$F_G = -\int_L \left(\sigma_{ij}^A + \sigma_{ij}^I\right) n_j b_i dx_3', \tag{5.89}$$

where n_j is a unit vector normal to the glide plane and the integration is over the line of the threading segment in its glide plane. Determination of this glide force is complex and depends on the shape of the threading dislocation. However, Freund considered a simplified treatment that avoids the need for knowing the dislocation shape.

Freund considered that in part of the glide plane, the force of the misfit dislocation negates the net glide force due to residual strain in the film. The glide force per unit length acting on an infinitesimal element of the threading arm is

$$w(x_2, x_3) = w^A + w^I(x_2, x_3) = \frac{2\pi(1-\nu)h}{Gb^2}\sigma_{ij}^A n_j b_i + \frac{2\pi(1-\nu)h}{Gb^2}\sigma_{ij}^I n_j b_i, \tag{5.90}$$

For an arbitrary value of the background strain ε*, this equation may be used to determine the region in which the glide force due to this strain is negated by the action of the misfit dislocation, and this corresponds to the shaded area in Figure 5.37. Therefore, the remaining channel height h^* through which the threading dislocation must pass may be determined by numerical solution of Equation 5.89 for the four combinations of Burgers

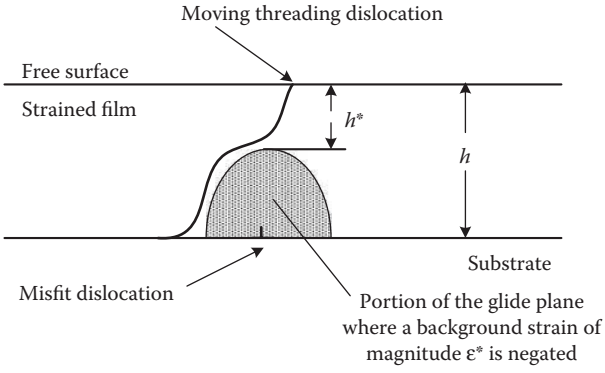


FIGURE 5.37
Gliding threading dislocation approaches a misfit dislocation in the interface. The glide force on the threading dislocation is negated in the shaded region, so that the threading dislocation must squeeze through the remaining thickness h^*. (Adapted from V. T. Gillard et al., *J. Appl. Phys.*, 76, 7280 [1994]. With permission. Copyright 1994, American Insitute of Physics.)

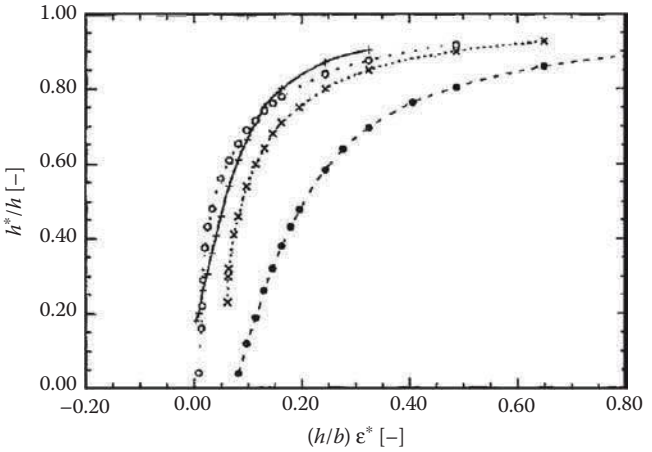


FIGURE 5.38
Normalized channel height for glide of a threading dislocation above an orthogonal misfit dislocation for a film containing a residual strain ε*. The four curves were calculated for the four independent combinations of Burgers vectors. (Adapted from V. T. Gillard et al., *J. Appl. Phys.*, 76, 7280 [1994]. With permission. Copyright 1994, American Institute of Physics.)

vectors, and these results are given in Figure 5.38. It can be seen that, for small values of the residual strain ε*, the force acting on the dislocation is negated over most of the film thickness so that h^*/h approaches zero. For large values of the residual strain, h^*/h approaches unity, so that h^* approaches h. Therefore, a dislocation that is pinned in a thin layer may become unpinned in a growing film as the layer thickness increases. On the other hand, a mobile dislocation may become pinned if the residual strain decreases during the growth or annealing of a layer.

Figure 5.39 compares experimental and theoretical values of the residual strain as a function of thickness for SiGe/Si(001). Experimental values were determined by measurements of the wafer curvature and by using Stoney's equation,[96] for limited reaction processing (LRP) layers of $Si_{0.655}Ge_{0.345}$/Si, chemical-vapor-deposited (CVD) $Si_{0.895}Ge_{0.105}$/Si, and MBE-grown SiGe/Si with 9%–26% germanium.[100,101] The theoretical curve labeled "isolated dislocation" was calculated without taking pinning into account. The remaining curves were calculated for the four combinations of Burgers vectors, and it can be seen that the one corresponding to the strongest blocking (parallel Burgers vectors) is in fairly close agreement with the experimental data.

5.6.4 Surface Roughness and Dislocation Blocking

In the previous discussion of dislocation blocking, the surface was assumed to be perfectly smooth. Samavedam and Fitzgerald[102] have shown that, in the presence of surface roughness, gliding threading dislocations may be blocked by the combined effects of orthogonal misfit dislocations and surface trenches, as shown in Figure 5.40. Here, the surface trench is sufficiently deep so that it overlaps the region where the net glide force on the gliding dislocation is cancelled by the strain field of the orthogonal misfit dislocation. Even in a less extreme case, such as the one shown in Figure 5.41, the surface roughness of amplitude r decreases the height of the channel in which the threading dislocation may glide past the orthogonal misfit dislocation to $h^* - r = h - R_0 - r$, where h is the average film thickness and R_0 is the blocking radius of the orthogonal misfit dislocation. As long as this channel

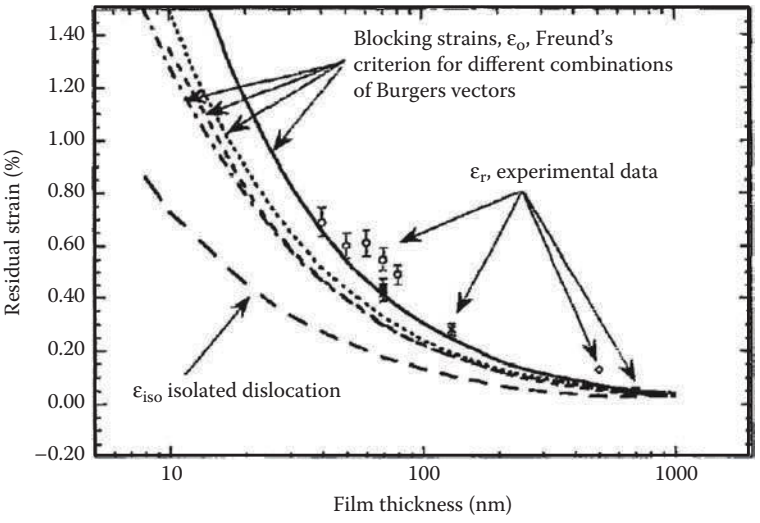


FIGURE 5.39
Residual strain as a function of thickness for SiGe/Si(001). (Adapted from V. T. Gillard et al., *J. Appl. Phys.*, 76, 7280 [1994]. With permission. Copyright 1994, American Institute of Physics.)

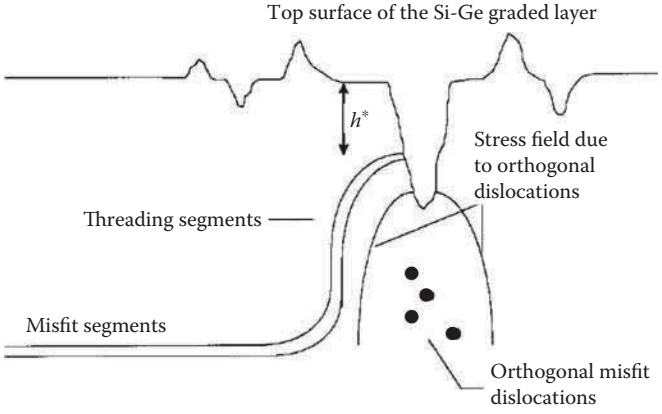


FIGURE 5.40
Gliding threading dislocations being blocked by the combined action of orthogonal misfit segments and surface roughness at the sidewall of a trench. (Reprinted from S. B. Samavedam and E. A. Fitzgerald, *J. Appl. Phys.*, 81, 3110 [1997]. With permission. Copyright 1997, American Institute of Physics.)

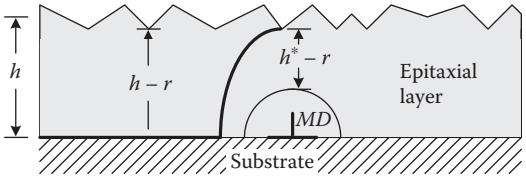


FIGURE 5.41
A gliding dislocation may squeeze through the reduced channel height $h^* - r$, where r is the amplitude of the surface roughness, as long as the glide force can overcome the line tension in this reduced thickness.

height is greater than the critical layer thickness, the gliding dislocation may pass by the orthogonal misfit dislocation. In this case, the glide force on the threading dislocation is able to overcome the line tension in the reduced channel height.

5.6.5 Matthews, Mader, and Light Kinetic Model

The first kinetic model for lattice relaxation was developed by Matthews, Mader, and Light.[45] As in the equilibrium model, they considered the forces acting on a grown-in threading dislocation. The glide force exerted on the dislocation, which tends to make it glide in a sense so as to produce a length of misfit dislocation in the interface, is

$$F_G = fYb\cos\lambda\cos\phi\,, \tag{5.91}$$

where:

- f is the lattice mismatch
- Y is the biaxial modulus
- b is the length of the Burgers vector
- λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface
- ϕ is the angle between the interface and the normal to the slip plane

In the isotropic case considered by Matthews, Mader, and Light, the line tension of the misfit segment of the dislocation is given by

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{5.92}$$

where:

- ν is the Poisson ratio
- α is the angle between the Burgers vector and the line vector for the dislocations
- G is the shear modulus (assumed to be the same for the epitaxial layer and the substrate)

For the gliding dislocation, there is also a Peierls force (lattice friction force) that opposes the motion. Following the work of Haasen,[103] Matthews, Mader, and Light assumed that the Peierls force acting on the bowing dislocation was given by

$$F_F = \left(\frac{h}{\cos\phi}\right)\left(\frac{vk_BT}{bD_0}\right)\exp\left(\frac{U}{k_BT}\right), \tag{5.93}$$

where:

- h is the layer thickness
- v is the dislocation glide velocity
- D_0 is the diffusion constant of the dislocation core
- U is the activation energy for the diffusion of the dislocation core
- k_B is the Boltzmann constant
- T is the absolute temperature

The linear density of misfit dislocations was considered to be constant with time (dislocation multiplication processes were not included). Within this assumption, the time rate of change of the lattice relaxation is

$$\frac{d\delta}{dt} = vDb\cos\lambda\,, \tag{5.94}$$

where D is the threading dislocation density in the substrate. Solving, in the anisotropic case, the time-dependent lattice relaxation is given by

$$\delta = \beta\left[1 - e^{-\alpha t}\right], \tag{5.95}$$

where*

$$\alpha = \frac{Gb^3D(1+\nu)\cos\phi\cos^2\lambda D_0 \exp\left(-\dfrac{U}{k_BT}\right)}{2(1-\nu)k_BT}, \tag{5.96}$$

and β is the limiting (equilibrium) value of the lattice relaxation for the layer,

$$\beta = \left|f - \varepsilon_{\|,\mathrm{eq}}(h)\right| = \left|f\right| - \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi h(1-\nu)\cos\lambda}\left[\ln\left(\frac{h}{b}\right)+1\right]. \tag{5.97}$$

Figure 5.42 shows the predicted behavior for heteroepitaxial Ge on GaAs(110). In the figure, the filled circles represent data for a sample with $\alpha t \ll 1$. The data for this sample fall on the dashed curve calculated for $\alpha t \ll 1$ (the curve corresponding to coherently strained material, marked A in the figure). The curve marked B in the figure corresponds to $\alpha t \gg 1$ (the equilibrium curve). The data point shown by the open circle was measured for a sample with $\alpha t \approx 1$, and closely matches the curve calculated for $\alpha t = 1$ (labeled C in the figure). Therefore, depending on the value of αt, it is possible to grow samples with equilibrium values of strain, or values that greatly exceed the predictions of equilibrium theory.

5.6.6 Dodson and Tsao Kinetic Model

Dodson and Tsao[104] built upon the Matthews, Mader, and Light model by including a dislocation multiplication term. In the Dodson and Tsao model, it was assumed that the glide velocity for a dislocation follows the empirical relationship

$$v = B\tau_{\mathrm{eff}}^{m}\exp\left(-\frac{U}{k_BT}\right), \tag{5.98}$$

* The equation given here differs by a factor of four from that given by Matthews, Mader, and Light, due to the correction of Fitzgerald (E. A. Fitzgerald, Mater. Sci. Rep., *7*, 87 [1991]).

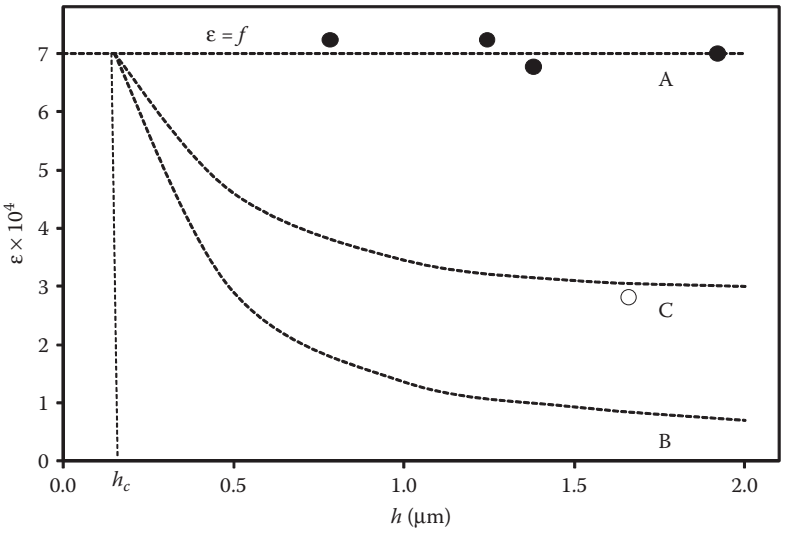


FIGURE 5.42
Elastic strain vs. thickness for heteroepitaxial Ge/GaAs(011). The filled circles represent data for a sample with $\alpha t \gg 1$, and the open circle was measured for a sample with $\alpha t \approx 1$. The dashed curves were calculated from the kinetic model. For curve A, it was assumed that $\alpha t \gg 1$, and for curve C, it was assumed that $\alpha t = 1$. (Adapted from J. W. Matthews et al., *J. Appl. Phys.*, 41, 3800 [1970]. With permission. Copyright 1970, American Institute of Physics.)

where:

τ_{eff} is the effective stress
B is a constant

(In most materials, the exponent m is found to be between 1 and 2; Dodson and Tsao assumed the value to be unity.)

The effective stress is the driving force for lattice relaxation. Dislocation multiplication was modeled by the assumption

$$\frac{d\rho}{dt} = K\rho_m v \tau_{\text{eff}} , \tag{5.99}$$

where:

ρ_m is the density of mobile dislocations
K is a phenomenological parameter

The strain relief is proportional to the linear density of misfit dislocations, so that

$$\frac{d\gamma}{dt} = b \sin\alpha \sin\phi \frac{d\rho}{dt} . \tag{5.100}$$

Combining Equations 5.68 and 5.69, with the assumption that all dislocations are mobile so that $\rho_m = \rho$, we obtain

$$\frac{d\gamma}{dt} = \kappa v |\tau_{\text{eff}}| \gamma(t) , \tag{5.101}$$

where κ is a constant. The absolute value of the effective stress in the heteroepitaxial layer is

$$\tau_{\text{eff}} = \frac{2G(1+\nu)}{(1-\nu)} \cos\lambda \cos\phi \left[|f| - |\gamma(t)| - |\varepsilon_{\parallel,\text{eq}}(h)| \right]. \tag{5.102}$$

Combining the above equations, we obtain a single differential equation for the time-dependent relaxation:

$$\frac{d\gamma(t)}{dt} = C\left[|f| - |\gamma(t)| - |\varepsilon_{\parallel,\text{eq}}(h)| \right]^2 \left(\gamma(t) + \gamma_0 \right), \tag{5.103}$$

where γ_0 represents the lattice relaxation from initial sources of dislocation in the substrate. The Dodson and Tsao model has been applied to a number of heteroepitaxial systems, in an attempt to better understand their lattice relaxation processes. This involves the adjustment of the parameters C and γ_0 to produce a good fit with the measured results. For example, Dodson and Tsao[101] applied this model to the case of SiGe/Si(001) grown by Bean et al.,[9] at a temperature of 823 K, or roughly 70% of the growth temperature. They found that the experimentally measured strains could be fit using $CG^2 = 46\ \text{s}^{-1}$ and $\gamma_0 = 3\times10^{-5}$. More recently, Yarlagadda et al.[105] applied the Dodson and Tsao model to $ZnSe_{1-x}Te_x$/InGaAs/InP (001), grown at 653 K, or roughly 40% of the melting temperature. In that work, it was necessary to use $CG^2 = 80\ \text{s}^{-1}$ and $\gamma_0 = 10^{-9}$ in order to reproduce the experimental results. The calculations are insensitive to the value of γ_0. It is significant, however, that Yarlagadda et al. needed to use a value of CG^2 that was roughly twice the value utilized by Dodson and Tsao. This indicates that dislocations glide more readily in $ZnSe_{1-x}Te_x$ at 40% of the growth temperature than in SiGe at 70% of the growth temperature.

Although the Dodson and Tsao model provides a reasonable description of the lattice relaxation process in a number of material systems, it suffers from three important liabilities that have limited its widespread use. First, the relaxation is given by a nonlinear hereditary integral equation, which does not lend itself to a simple analytical solution, but requires numerical solution.[106] Second, the empirical parameters used in the model have been determined for only a few material systems of interest. Third, the model does not lend itself readily to inclusion of the temperature dependence of the glide velocity or the stress dependence of the activation energy for glide, because these are included implicitly in the fitting parameter C. Despite these limitations, application of the Dodson and Tsao model has served as a starting point for understanding lattice relaxation in a number of material systems.

5.6.7 Hull, Bean, and Buescher Kinetic Model

Hull, Bean, and Buescher[83,84] developed a model for strain relaxation in $Si_{1-x}Ge_x$/Si(001) heterostructures that included a stress-dependent activation energy for glide of dislocations. This model is based on the idea that the total misfit dislocation length per unit area $L(t)$, which determines the extent of lattice relaxation, is equal to the sum of the lengths of the individual misfit dislocations:

$$L(t) = \int_0^t N(t')v(t')dt', \tag{5.104}$$

where:

$N(t')$ is the number of growing misfit dislocations per unit area
$V(t')$ is the glide velocity of threading arms

In a growing film, the lower limit of integration corresponds to the point in time when the critical layer thickness is exceeded. The glide velocity $v(t')$ was assumed to be

$$v(t') = v_0 \exp\left(-\frac{E_a(\sigma)}{kT}\right)\tau_{\text{eff}}, \tag{5.105}$$

where:

$v_0 = 3.3\times10^7\,\text{cm}^3\text{s}^{-1}\text{dyn}^{-1}$
τ_{eff} is the effective stress in dyn/cm^2
k is the Boltzmann constant
T is the absolute temperature
$E_a(\sigma)$ is the stress-dependent activation energy for glide

assumed to be given by

$$E_a(\sigma) = E_a(0)\left[1-\frac{\tau_{\text{eff}}}{\sigma_c}\right], \tag{5.106}$$

where:

$E_a(0)$ is a constant, assumed to be 2.0 eV for $Si_{0.75}Ge_{0.25}$
σ_c is the critical stress, determined to be $1.5\times10^{10}\,\text{dyn/cm}^2$ for $Si_{0.75}Ge_{0.25}$

The number of growing dislocations per unit area was assumed to be

$$N(t') = N_0 + N_1 t' \exp\left(-\frac{E_n}{kT}\right) + k_s - N_p P_m N_i - P_p N_i, \tag{5.107}$$

where:

N_0 is the density of preexisting dislocations
$N_1 t' \exp(-(E_n/kT))$ represents the "thermally nucleated dislocations"
$k_s P_m N_i$ represents dislocation generation by multiplication mechanisms

where:

N_i is the number of dislocation interactions per unit area
P_m is the probability of occurrence of a multiplication event
k_s represents the number of dislocations created by each multiplication event

A compelling feature of this model is that it predicts, at least in a qualitative way, the reduction in threading dislocation densities that has been observed in low-temperature growth of metamorphic buffers.[107] However, it cannot predict the reduction of threading dislocation density with film thickness, which is important in thicker layers, because it does not account for dislocation annihilation or coalescence. The term $P_p N_i$ represents the loss of mobile dislocations by pinning mechanisms, where P_p is the pinning probability. The areal density of dislocation interactions is given by

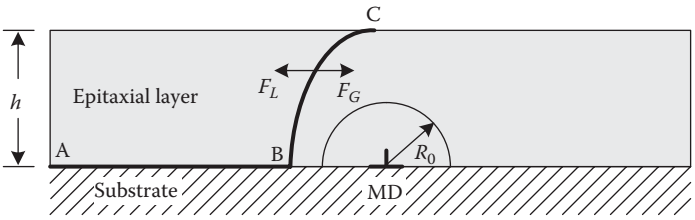


FIGURE 5.43
Glide of a threading dislocation toward an orthogonal misfit dislocation in the interface.

$$N_i = \frac{\pi^3 L^2(t)}{32}. \tag{5.108}$$

In order to calculate the pinning probability, Hull, Bean, and Buescher considered the glide of a threading dislocation toward an orthogonal misfit dislocation, as shown in Figure 5.43. Dislocation ABC contains a misfit segment AB and a threading segment BC. The threading dislocation segment BC is gliding to the right under the influence of a glide force F_G, which is opposed by the line tension in the misfit segment, F_L. If the orthogonal misfit dislocation labeled MD has a Burgers vector parallel to the gliding dislocation, there will be a repulsive force that cancels the net glide force $F_G - F_L$ within a "pinning radius" R_0.

In order to estimate the pinning radius Hull, Bean, and Buescher considered the repulsive force per unit length for two screw dislocations having parallel Burgers vectors and separated by a distance R, which is

$$\frac{F_{dp}}{L} = \frac{Gb^2}{2\pi R}. \tag{5.109}$$

Based on this relationship, the pinning radius for which the net force $F_G - F_L$ is cancelled is approximately

$$R_0 = \frac{Gb^2\sqrt{2}h}{2\pi\left(F_G - F_L\right)}, \tag{5.110}$$

where

$$F_G = \frac{2S(1+\nu)G\varepsilon bh}{(1-\nu)} \tag{5.111}$$

and

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\ln\left(\frac{4h}{b}\right). \tag{5.112}$$

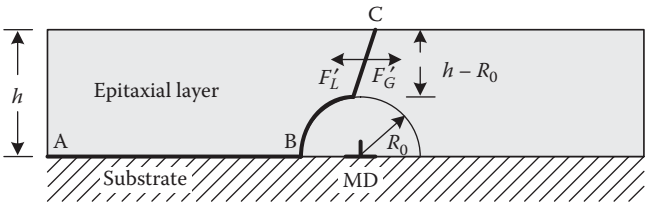


FIGURE 5.44
Pinning of a dislocation by an orthogonal misfit dislocation.

Now, to find if the gliding dislocation could still glide past the orthogonal misfit dislocation, Hull, Bean, and Buescher calculated the modified forces F'_G and F'_L as shown in Figure 5.44, given by

$$F'_G = \frac{2S(1+\nu)G\varepsilon b(h-R_0)}{(1-\nu)} \tag{5.113}$$

and

$$F'_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\ln\left(\frac{4(h-R_0)}{b}\right). \tag{5.114}$$

If $F'_G > F'_L$, then the dislocation would still be able to squeeze through the film thickness $h - R_0$, but otherwise, the dislocation would be considered to be pinned over its entire length. In such a case, the pinning probability was considered to be 1/9, which they considered the probability of the occurrence of parallel Burgers vectors. Furthermore, if the number of pinned dislocations was predicted to be such that $N(t')$ would be rendered negative, then it was assumed that only the thermally generated dislocations would be mobile, so that

$$N(t') = N_1 \Delta t \exp\left(-\frac{E_n}{kT}\right), \tag{5.115}$$

and the distance these dislocations would glide in time Δt would be the minimum of $v\Delta t$ or $9p$, where p is the average separation between misfit dislocations.

Using the equations above, Hull and Bean calculated the areal density of growing misfit dislocations $N(t')$, the total misfit dislocation length per unit area $L(t')$, the average length of misfit segments, $\bar{l} = L(t')/N(t')$, and the average separation of misfit dislocation segments, $p = 2/L(t')$. They found good agreement between calculations and experimental results for thin layers of $Si_{0.75}Ge_{0.25}$/Si(001) when dislocation pinning interactions were included, while assuming $N_m = 0$, as shown in Figure 5.45.

Three important differences between this model and the Dodson and Tsao model are that (1) Dodson and Tsao included dislocation multiplication, whereas Hull, Bean, and Buescher did not; (2) Hull, Bean, and Buescher included pinning, whereas Dodson and Tsao did not; and (3) Hull, Bean, and Buescher included a stress-dependent activation energy. Whereas both models seem consistent with experimental results, Dodson and Tsao considered much thicker layers, and it is not clear if both models can predict the behavior over the entire

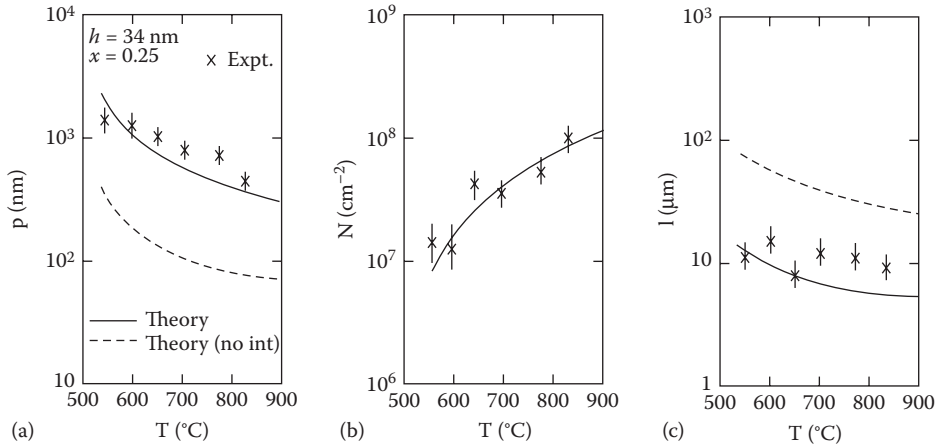


FIGURE 5.45
The average misfit dislocation spacing p, the areal density of threading dislocations N, and the average length of misfit dislocations $\bar{l}$ for 34 nm of $Si_{0.75}Ge_{0.25}$/Si(001) that was grown at 550°C at a rate of 0.3 nm/s and then annealed sequentially. The data point for 550°C corresponds to a 240 s anneal at 550°C. The 600°C data point corresponds to 240 s anneal at 550°C plus a 240 s anneal at 600°C, and so on. The solid lines were calculated with the inclusion of dislocation pinning effects, while pinning was ignored for the calculation of the dashed curves. (Reprinted from R. Hull et al., *J. Appl. Phys.*, 66, 5837 [1989]. With permission. Copyright 1989, American Institute of Physics.)

range of experimentally relevant thicknesses. It should also be noted that neither model includes annihilation or coalescence reactions between threading dislocations, and therefore neither can predict the reduction in the threading dislocation density in thicker layers.

5.6.8 Kujofsa et al.[108] Kinetic Model

Kujofsa et al.[108] extended the Dodson and Tsao[104] model by accounting for the time evolution of the equilibrium configuration and the inclusion of dislocation interactions. In contrast to the Dodson and Tsao model where the equilibrium configuration was determined for the fully grown film and the annealing was considered to be done postgrowth, with the full thickness, Kujofsa et al. approximated a structure with an arbitrary compositional profile by a stack of N lamina (sublayers) and considered each sublayer to be deposited one at a time and subsequently annealed for its growth period. After the deposition of a layer of the on-growing structure, the overall equilibrium configuration was determined in the partially grown structure and, from this, the rate of lattice relaxation as a function of time and distance from the interface (or each sublayer in the stack). Thus, the evolution of the equilibrium strain profile was considered. For a uniform epitaxial layer, the equilibrium configuration of the on-growing structure may be determined in a fashion similar to that described by Matthews. The lattice relaxation rate at a distance z above the interface of a mismatched epilayer is determined by the glide of dislocations in the underlying material. At abrupt interfaces, as in the case of a uniform layer, the linear misfit dislocation ρ_l is localized at the interface, giving rise to delta functions, and therefore the lattice relaxation rate at a distance z is determined by

$$\frac{d\gamma(z)}{dt} = KBb\sin\alpha\cos\lambda\tau_{eff}^2(z)\exp\left(-\frac{U}{kT}\right)\left[\rho_l(0)+\rho_0\right]. \tag{5.116}$$

where:

K and B	are constants
b	is the length of the Burgers vector
α	is the angle between the Burgers vector and line vector
λ	is the angle between the Burgers vector and the direction in the interface that is perpendicular to the intersection of the glide plane and the interface
τ_{eff}	is the effective stress that is determined by the difference of actual and equilibrium strain
U	is the activation energy for dislocation glide
k	is the Boltzmann constant
T	is the temperature
$\rho_l(0)$	is the linear density of misfit dislocations at the substrate interface
ρ_0	is a constant that represents the initial sources of dislocations

The effective stress is determined by the difference of the actual and equilibrium strain profiles in the material above and is given by

$$\tau_{\text{eff}}(z) = \left[\frac{2\cos\phi\cos\lambda}{h-z}\right]\int_z^h \left\{\frac{\mu(1-\nu)\left[\varepsilon_{\parallel} - \varepsilon_{\parallel,\text{eq}}\right]}{(1+\nu)}\right\} d\zeta, \tag{5.117}$$

where:

h	is the layer thickness
ϕ	is the angle between the surface normal and the slip plane
μ	is the shear modulus
ν	is the Poisson ratio
$\varepsilon_{\parallel}$	is the in-plane strain
$\varepsilon_{\parallel,\text{eq}}$	is the equilibrium in-plane strain

If the sample temperature is varied, either during growth or the postgrowth cooldown, the thermal strains introduced by the difference in thermal expansion coefficients are accounted for according to

$$\varepsilon_{\parallel}(T) = \varepsilon_{\parallel}(T_0) + \int_{T_0}^{T} (\alpha_s - \alpha_e)\, d\tau, \tag{5.118}$$

where $\varepsilon_{\parallel}(T_0)$ and $\varepsilon_{\parallel}(T)$ are the in-plane strains at temperatures T_0 and T, respectively, and α_s and α_e are the thermal coefficients of linear expansion for the substrate and epitaxial material, respectively. (Thermal strain is discussed in Section 5.8.1.) Kujofsa et al. applied this model to ZnSe/GaAs and ungraded ZnS_ySe_{1-y}/GaAs(001) to study the thickness and mismatch dependence of the in-plane strain.

5.6.9 Kinetically Limited Lattice Relaxation in Zinc Blende Semiconductors

The general consensus among the published body of experimental work is that the relaxation of semiconductor heterostructures exhibits four distinct thickness regimes with different relaxation rates. In the initial phase, the epilayer is coherently strained to the substrate and the relaxation rate remains zero until the critical layer thickness is reached.

In the secondary phase, the sluggish relaxation is governed by the slow formation of misfit dislocations from glide of preexisting threading dislocations (substrate-associated defects). In the third phase, relaxation occurs much more rapidly, whereby the increase in the effective stress drives most of strain relief from the continuous buildup of dislocations (nucleation and multiplication). In the last phase, the remaining strain relaxation proceeds more slowly (saturates) due to the reduced stress from the dislocation-mediated relaxation phase and possibly pinning of dislocations. This qualitative behavior is well described by assuming the thermally activated glide of dislocations. The observation of all four regimes has been reported in several material systems, including InGaAs/GaAs, ZnSe/GaAs, ZnSSe/GaAs, InAlAs/GaAs, InAlAs/InP, SiGe/Si, and InGaP/GaAs. In addition, the rate at which lattice relaxation takes place is governed by the activation energy for glide of each material system and the growth parameters.

5.6.9.1 Lattice Mismatch and Thickness Dependence of Kinetically Limited In-Plane Strain

Kujofsa et al.[109] investigated the in-plane strain as a function of epilayer thickness for ZnSe grown on GaAs(001) substrate by a two-step photoassisted MOVPE process in which the first 40% of the total thickness was grown at 595°C without photoassistance and the remaining 60% of the total thickness was grown at 360°C by photoassisted MOVPE. The kinetically limited strain relaxation (Figure 5.46) in these structures occurs much more gradually than predicted by the Matthews and Blakeslee equilibrium theory. The onset of lattice relaxation occurs at a thickness much greater than the critical layer thickness predicted by the equilibrium model. In addition, at a thickness of 800 nm, the epilayer exhibits ~60% relaxation, whereas equilibrium predicts ~90% relaxation.

5.6.9.2 Temperature-Graded Heterostructures

A drawback of uniform composition structures is that they allow limited flexibility in controlling the residual strain. From a fabrication point of view, the heterostructure is grown at an optimal temperature and growth rate with respect to the surface morphology or electronic properties, and the thickness then determines the in-plane strain and threading dislocation density. In some cases, the temperature may be varied, or graded, during

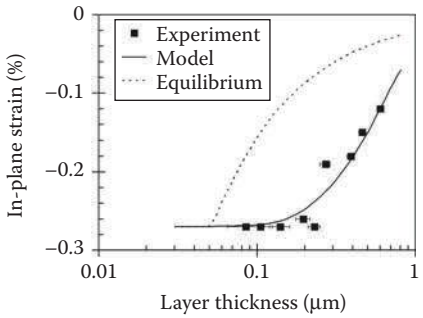


FIGURE 5.46
Room temperature in-plane strain as a function of epilayer thickness for ZnSe/GaAs(001) layers grown by a two-step MOVPE process. (Reprinted from T. Kujofsa et al., *J. Electron. Mater.*, 41, 2993 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

growth to afford greater control on the lattice relaxation and defects. This approach could also be applied in compositionally graded material.

Kujofsa and Ayers[110] performed a modeling study of temperature-graded ZnSe buffer layers to investigate this approach. The use of various temperature profiles allowed control of the thermal budget and influenced the final strain and defect densities. Heterostructures with convex-up (type A), linear (type B), and convex-down (type C) temperature grading profiles were considered. For type A structures, the first half of the total epilayer thickness was grown at constant temperature T_0, and in the remaining half of the total thickness, the temperature was linearly decreased from T_0 to a final temperature $T_F = 300°C$. In linear type grading, the temperature was varied linearly from T_0 to $T_F = 300°C$ during the epitaxial growth of the structure. For type C structures, in the growth of the first 50% of total epilayer thickness, the temperature was linearly decreased from T_0 to $T_F = 300°C$ and the remaining growth was continued at this same temperature.

Figure 5.47 shows the misfit (lattice mismatch), equilibrium strain, and actual strain as a function of the on-growing epilayer thickness. The results are suggestive of the four-regime lattice relaxation behavior described previously. Approximately two-thirds of the mismatch strain is relaxed. The structure with the convex-up temperature profile, and therefore a higher thermal budget, exhibits a greater extent of lattice relaxation and therefore lower strain values, as shown in Figure 5.47a. The structure with the convex-down temperature profile, and therefore the minimum thermal budget, exhibits very little strain relaxation and high residual strain, as shown in. However, it can also be seen from the results of Figure 5.47c. The structure grown with a linear temperature profile is intermediate, with greater strain values than for the convex-down structure, but the strain relaxation is still suggestive of the four-regime behavior.

The effect of changing the initial temperature was also investigated in this study. Figure 5.48 shows the average in-plane strain as a function of the epilayer thickness with initial temperature as a parameter for type A (Figure 5.48a), type B (Figure 5.48b), and type C (Figure 5.48c) structures. The initial temperature ranged from 400°C to 600°C in steps of 50°C, and the final temperature was fixed at $T_F = 300°C$. The average in-plane strain decreases monotonically with increasing thickness, and a higher initial temperature increases the overall thermal budget and therefore the total amount of relaxation. Furthermore, an increase in the available thermal budget (higher integrated temperature profile) enhances the strain relaxation rates; this is evident by the complete state of relaxation at a particular temperature or by the degree of change in the in-plane strain for increasing thickness. For example, at $T_0 = 600°C$, structures with type A temperature grading profile exhibit near complete strain relaxation at a thickness of ~0.75 μm, whereas for type B and C, near complete relaxation does not occur until a thickness of ~1.35 μm and $h \sim 2.6$ μm, respectively. Therefore, for a given temperature grading profile, there exists a combination of starting temperature and layer thickness that results in a given residual strain. Therefore, the residual strain may be tailored independently of the thickness and the final growth temperature.

5.6.9.3 Lattice Relaxation in the InGaAs/GaAs Material System

Although the kinetically limited lattice relaxation of $In_xGa_{1-x}As/GaAs(001)$ is consistent with the four-thickness-regime behavior described previously, a number of workers have reported that dislocation blocking[111] is especially important in the InGaAs/GaAs material system,[112–115] and that the strain may saturate at a nonequilibrium value. Several groups have focused on the detailed study of lattice relaxation by *in situ* or postgrowth methods.

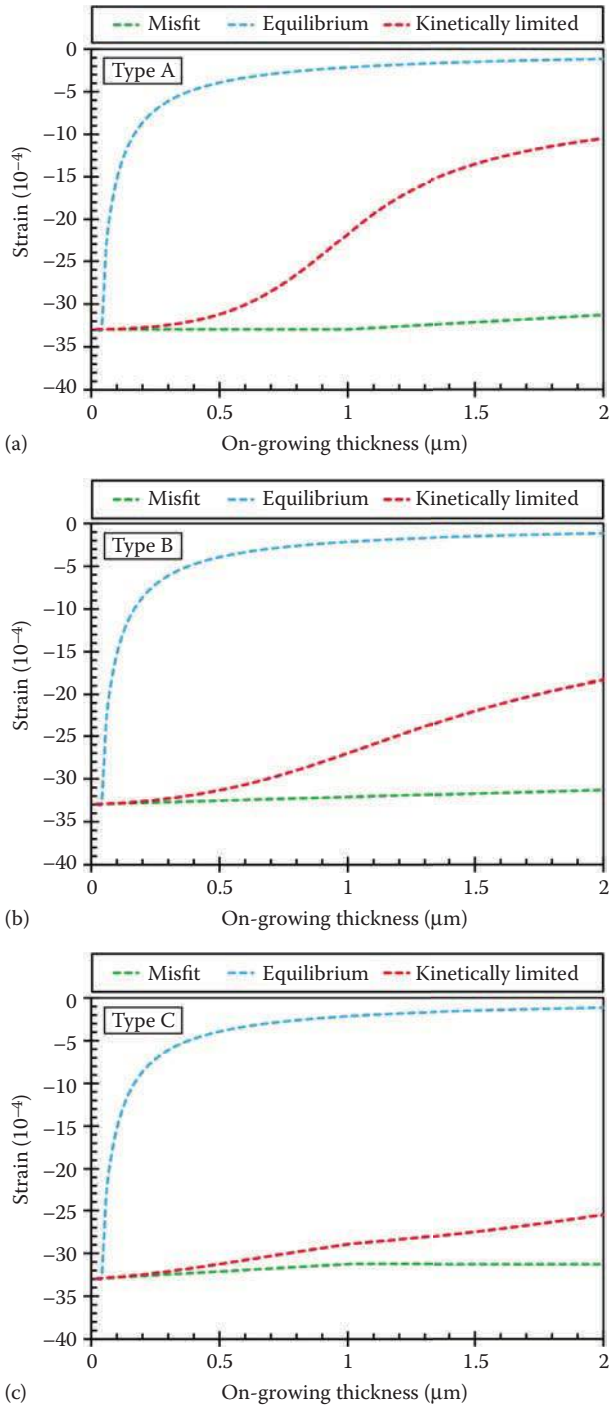


FIGURE 5.47
The evolution of strain (kinetically limited, misfit, and equilibrium) as a function of the grown thickness for type (a) A, (b) B, and (c) C structures. The initial T_0 and final T_F temperatures in these structures were 400°C and 300°C, respectively. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 44, 3030 [2015]. With permission. Copyright 2015, Minerals, Metals and Materials Society.)

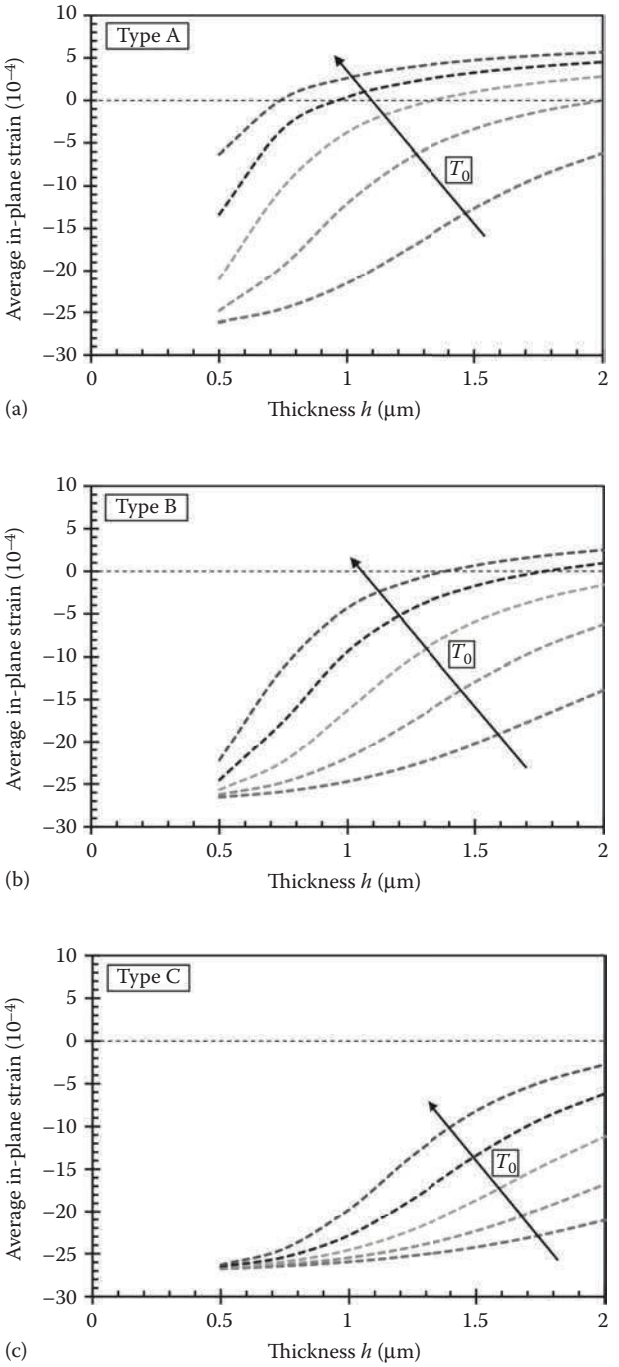


FIGURE 5.48
Average kinetically limited in-plane strain as a function of the final epilayer thickness h with initial temperature as a parameter for type (a) A, (b) B, and (c) C structures. The initial temperature T_0 is varied from 400°C to 600°C in steps of 50°C, and the final T_F temperature is fixed at 300°C. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 44, 3030 [2015]. With permission. Copyright 2015, Minerals, Metals and Materials Society.)

Whaley and Cohen[116] studied the strain relaxation of MBE-grown InGaAs/GaAs *in situ* by reflection high-energy electron diffraction (RHEED). The temperature dependence of the lattice relaxation was investigated in a temperature range low enough so as to prevent indium evaporation. In the initial stages of lattice relaxation, it was found that the relaxation behavior was strongly temperature dependent, as would be expected for thermally activated glide. This can be seen in the results in Figure 5.49a, which shows the relaxation of layers with 33% indium at four different growth temperatures. The initial relaxation becomes progressively more sluggish as the growth temperature is decreased from 510°C to 470°C, so that the apparent critical layer thickness increases with decreasing temperature. For growth at 470°C, the onset of measurable relaxation appears to occur at a thickness greatly exceeding the Matthews and Blakeslee critical layer thickness. Figure 5.49b shows the relaxation of InGaAs layers having different indium mole fractions, but all grown at a fixed temperature of 510°C. Clearly, the apparent critical layer thickness varies strongly with the indium mole fraction, but it does not follow the reciprocal relationship predicted by Matthews and Blakeslee, showing the importance of kinetic limitations in the initial relaxation.

A fascinating finding from Whaley and Cohen's work was that the lattice relaxation ceased during growth interruptions, as shown in Figure 5.50. This figure shows the evolution of the in-plane lattice parameter determined by RHEED as a function of time for two deposited $In_xGa_{1-x}As$ films with total thicknesses of 50 and 75 nm. The nominal indium composition in both films was 16% and the growth temperature was 450°C. The calculated Matthews and Blakeslee critical layer thickness for this composition is 4.6 nm, whereas the onset of lattice relaxation was observable at ~15 nm for these structures. The overall lattice relaxation behavior of both samples could be fit approximately by the Dodson and Tsao model up to the point in time at which the growth is interrupted, at 250 s for the 50 nm film and 375 s for the 75 nm sample. Qualitatively, this could be explained by dislocation blocking interactions. Pinned dislocations can become unpinned by the growth of additional material, but if the growth is interrupted, then this process stops and it is possible for all of the dislocations to exist in a pinned state. Although this is a likely explanation, it should be mentioned that these measurements revealed only the surface in-plane lattice constant, which could have been different from the value at the substrate interface. In addition, Waley and Cohen needed to use a large value of the activation energy for glide (4.4 eV), much larger than the bulk values given by Yonenaga for InAs[117] and GaAs.[118] This could have been in part because the equilibrium strain was assumed constant, and its thickness variation was not accounted for in applying the Dodson and Tsao relationship.

Beresford et al.[119] (see Figure 5.51) and Sasaki et al.[120] studied the evolution of the residual strain in InGaAs/GaAs(001) using an *in situ* multibeam optical stress sensor (MOSS) and *in situ* x-ray reciprocal space mapping (RSM). They experimentally verified the four thickness regimes with varying relaxation rates, but they were unable to make any conclusive statements on the major role of thermally activated glide for lattice relaxation. However, dislocation flow calculations from this work suggested that a single fitting parameter BK (see Equation 5.71) was not adequate to fit the experimental results over the thickness range considered. Beresford argued that in relatively thick structures, the BK value is insensitive to the growth temperature, and therefore the behavior is not consistent with thermally activated glide described by a single activation energy. This could be the result of a stress-dependent activation energy, as predicted for the case of the diffusion kink model of glide, or it could be related to dislocation blocking, which is relatively insensitive to temperature. The work of Dunstan[121] suggests that the suppression of glide velocity and approximate temperature independence occurs as a result of dislocation blocking or

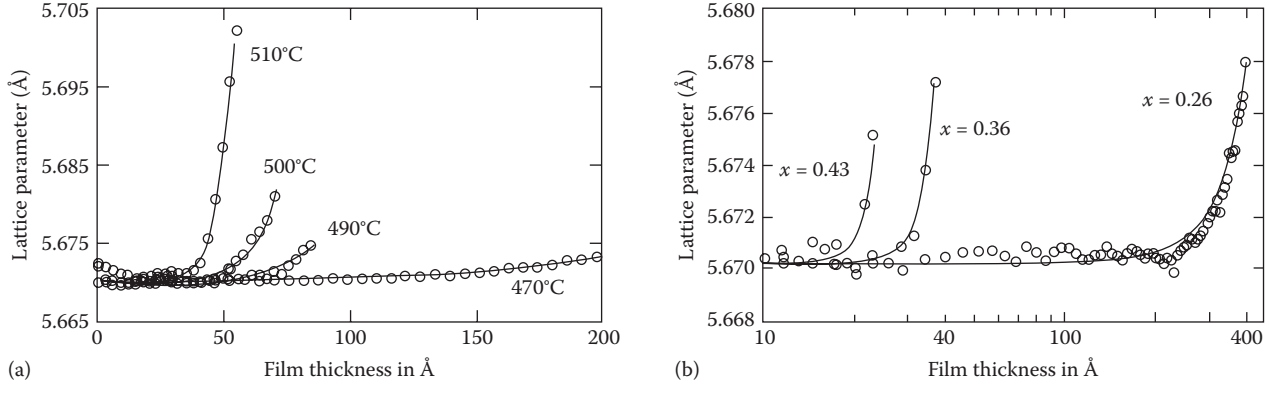


FIGURE 5.49
Measured in-plane lattice parameter of $In_xGa_{1-x}As/GaAs$ grown on a GaAs(001) substrate. (a) The epilayer contains an indium composition of 33% ($x = 0.33$) and deposited at various growth temperatures. (b) The growth temperature is fixed at 510°C and the epilayer contains various indium compositions. The symbols are experimental results, whereas solid lines represent fits to the Dodson–Tsao model. (Reprinted from G. J. Whaley and P. I. Cohen, *Appl. Phys. Lett.*, 57, 144, 1990. With permission. Copyright 1990, American Institute of Physics.)

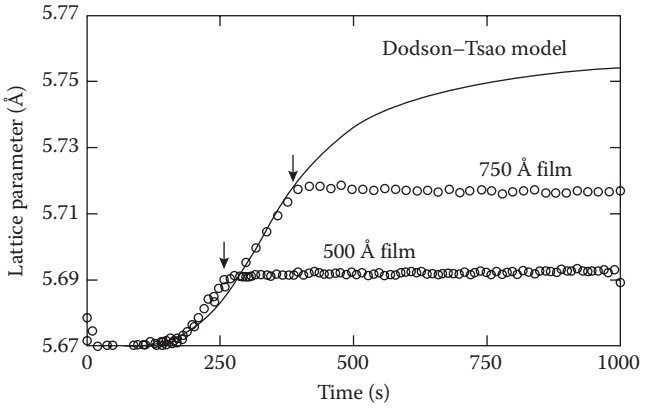


FIGURE 5.50
Measured in-plane lattice parameter of $In_xGa_{1-x}As$ films deposited on a GaAs(001) substrate as a function of the growth time for various thicknesses (50 and 75 nm). The indium composition was fixed at 26% ($x = 0.26$) and the growth temperature was 470°C. The solid line is a Dodson–Tsao fit to the 75 nm film. The arrows show the point at which growth was interrupted. (Reprinted from G. J. Whaley and P. I. Cohen, *Appl. Phys. Lett.*, 57, 144 [1990]. With permission. Copyright 1990, American Institute of Physics.)

the onset of dislocation climb. In the work of Sasaki et al.,[120] different values of BK were used for each of the four relaxation regimes in order to describe the experimental results. This approach is phenomenological, however, and provides little insight into the physical processes involved.

Lynch et al.[122] investigated the temperature dependence of strain relaxation in MBE-grown $In_{0.14}Ga_{0.86}As$ on GaAs and noticed that an increase in temperature from 443°C to 469°C led to the formation of a new regime with enhanced relaxation rates and which commences at a thickness below that of the rapid relaxation phase. They referred to this as the shoulder regime (shown in Figure 5.52) and noticed that about ~17% of the initial misfit stress was relaxed in this regime. To better understand the mechanics for this relaxation regime, they studied the dislocation structure of samples before and after the shoulder regime. Their results indicated that there is a significant increase in the dislocation density after this regime, with higher densities in the [110] direction,[123] and they also noted the presence of edge dislocations. From multibeam optical stress sensor measurements, they concluded that the relaxation in this regime was governed by multiplication rather than nucleation of dislocations. Based on this, it was further concluded that the dislocation multiplication is strongly temperature dependent based on the observation that the shoulder in relaxation is suppressed at temperatures below 443°C.

Despite the existence of a large body of experimental data in the InGaAs/GaAs material system, the underlying physics remains poorly understood, and this has led to the widespread use of phenomenological models to describe the behavior. Along these lines, two types of phenomenological behavior have commonly been reported: in the first, the in-plane strain is considered to be inversely proportional to thickness[124,125] $\varepsilon_{\|} \propto h^{-1}$, and in the second, the in-plane strain is considered to follow an inverse square root dependence[110,126] given by $\varepsilon_{\|} \propto h^{-1/2}$. Krishnamoorthy et al.[127] developed yet another phenomenological model that was piecewise-linear with two regimes. This model was used to describe measured residual strains in MBE-grown InGaAs on GaAs with varying composition and thickness, as determined by high-resolution x-ray diffraction. The results of Figure 5.53 show that the residual strain is a strong function of both the epilayer thickness and composition, and

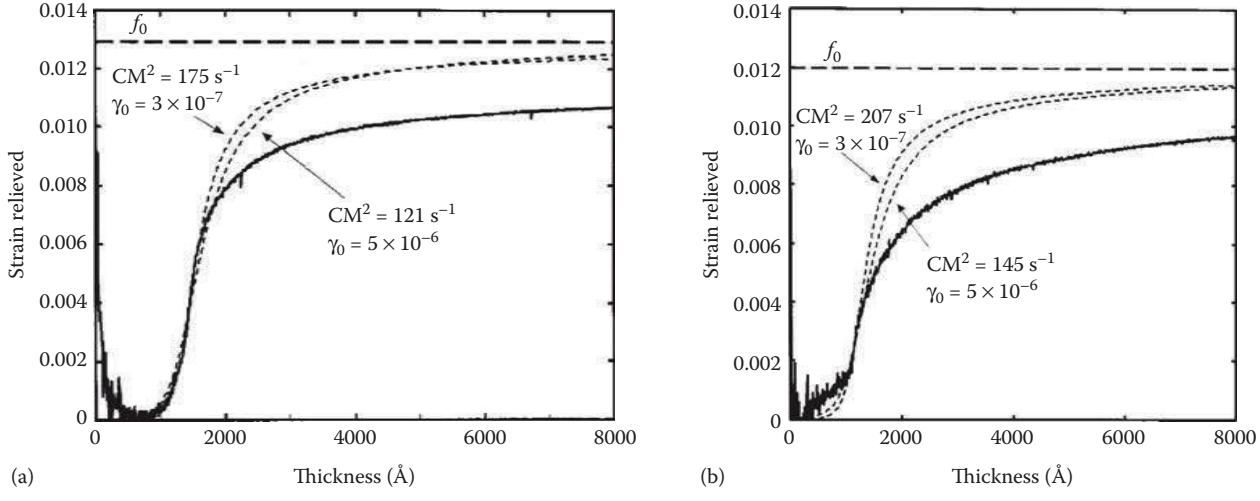


FIGURE 5.51
Strain relaxation as a function of the on-growing thickness for epilayers with (a) a nominal indium composition of 18% and a growth temperature of 443°C and (b) a nominal indium composition of 17% and a growth temperature of 522°C. (Reprinted from R. Beresford et al., *J. Vac. Sci. Technol. B*, 18, 1431 [2000]. With permission. Copyright 2000, American Vacuum Society.)

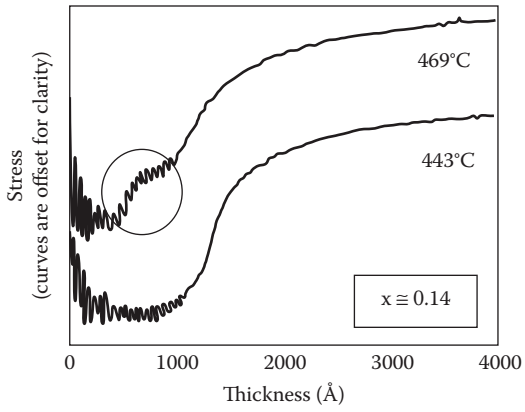


FIGURE 5.52
Stress evolution of In0.14Ga0.86As on GaAs substrates. (Reprinted from C. Lynch et al., *J. Vac. Sci. Technol. B*, 20, 1247 [2002]. With permission. Copyright 2002, American Vacuum Society.)

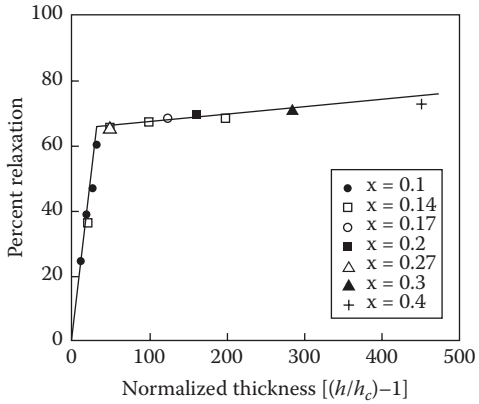


FIGURE 5.53
Relaxation percentage as a function of normalized thickness for InGaAs on GaAs(001) heterostructures with various epilayer thickness and composition. (Reprinted from V. Krishnamoorthy et al., *Appl. Phys. Lett.*, 61, 2680 [1992]. With permission. Copyright 1992, American Institute of Physics.)

were fit with a piecewise linear model as follows. In the initial stages of growth, relaxation occurs rapidly and the percentage relaxation was modeled by

$$\%R = (2.0 \pm 0.1)\left(\frac{h}{h_c} - 1\right);\ h_c \le h \le 30h_c\ , \tag{5.119}$$

whereas after the rapid relaxation phase, it was reported that there was a saturation region that could be modeled by

$$\%R = (0.020 \pm 0.005)\left(\frac{h}{h_c} - 1\right) + (66 \pm 1);\ 30h_c < h \le \sim 450h_c\ . \tag{5.120}$$

Dunstan et al.[121] applied a reciprocal of thickness model for InGaAs on GaAs structures and concluded that the strain thickess product in the rapid relaxation regime is a constant, described by

$$\varepsilon_{\|} h = (0.8 \pm 0.1)\ \text{nm}. \tag{5.121}$$

This expression was reported to be valid for thicknesses from the critical layer thickness up to approximately 10 times the critical layer thickness (the onset of the so-called work-hardening regime).

5.6.10 Kinetically Limited Relaxation in Nitride Semiconductors

There have been few fundamental studies of kinetically limited strain relaxation in nitride semiconductors, and the dislocation dynamics are poorly understood. It is known that the lattice relaxation behavior differs between planar nitride layers and ELO material, apparently because of the difference in the threading dislocation density.

Song et al.[128] studied the kinetically limited strain relaxation in $In_{0.1}Ga_{0.9}N$ layers of varying thickness grown on GaN by MOVPE. *c*-Plane sapphire substrates were used, on which were grown a 25 nm low-temperature GaN nucleation layer, followed by a thick 1 μm layer of GaN grown at 1000°C. $In_{0.1}Ga_{0.9}N$ layers of varying thickness from 100 to 300 nm were then grown on the thick GaN at 750°C, and the final state of strain for these layers was determined by high-resolution x-ray diffraction using the 00.2 reflection.

The 00.2 x-ray rocking curves for the five samples of varying $In_{0.1}Ga_{0.9}N$ thickness are shown in Figure 5.54. By fitting the GaN and InGaN diffraction peaks to Lorentzian functions, the centers of the peaks were accurately determined and used to find the in-plane and out-of-plane strains in the $In_{0.1}Ga_{0.9}N$ layers, shown in Figures 5.55 and 5.56, respectively. These results show that the strains are constant up to the critical layer thickness of ~200 nm. Beyond that, both strains decrease hyperbolically with thickness:

$$\varepsilon_{\|} = \frac{K_{P0}}{h} \tag{5.122}$$

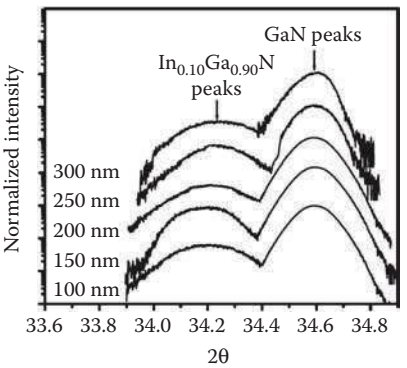


FIGURE 5.54
00.2 x-ray rocking curves for $In_{0.1}Ga_{0.9}N$/GaN/sapphire samples with varying $In_{0.1}Ga_{0.9}N$ thickness. (Reprinted from T. L. Song et al., *Adv. Mater. Micro Nano Syst.* [2004]. With permission. Copyright 2004, Dspace@mit.)

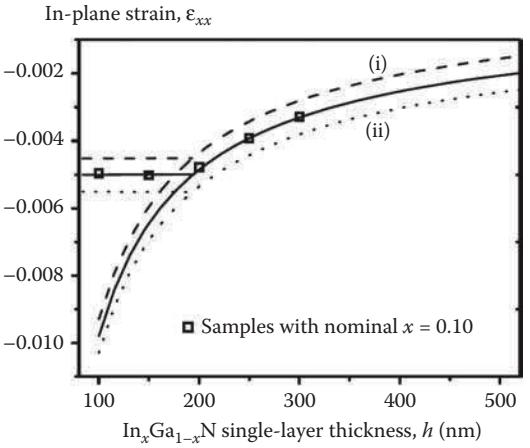


FIGURE 5.55
In-plane strain as a function of $In_{0.1}Ga_{0.9}N$ thickness. $In_{0.1}Ga_{0.9}N$/GaN/sapphire samples. (Reprinted from T. L. Song et al., *Adv. Mater. Micro Nano Syst.* [2004]. With permission. Copyright 2004, Dspace@mit.)

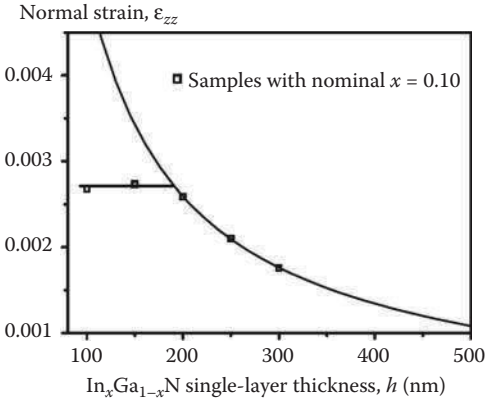


FIGURE 5.56
Out-of-plane strain as a function of $In_{0.1}Ga_{0.9}N$ thickness. In0.1Ga0.9N/GaN/sapphire samples. (Reprinted from T. L. Song et al., *Adv. Mater. Micro Nano Syst*. [2004]. With permission. Copyright 2004, Dspace@mit.)

and

$$e_{\perp} = \frac{K_{N0}}{h}, \tag{5.123}$$

where $K_{P0} = -0.98 \pm 0.03\,\text{nm}$ and $K_{N0} = +0.51 \pm 0.03\,\text{nm}$. These reciprocal relationships agree qualitatively with the Matthews and Blakeslee force balance equilibrium model. Based on their observations, Song et al. speculate that other compositions of InGaN, and indeed other nitride semiconductors, will exhibit similar reciprocal behavior, although the underlying mechanisms are incompletely understood.

5.7 Lattice Relaxation on Vicinal Substrates: Crystallographic Tilting of Heteroepitaxial Layers

Heteroepitaxial semiconductors grown on vicinal substrates generally exhibit a crystallographic tilt with respect to the underlying substrate. Thus, for nominal (001) heteroepitaxy of a zinc blende semiconductor, the [001] axes for the deposit and substrate are not parallel if the substrate [001] axis is inclined from the normal. This effect has been observed in many material systems, including GaN/Al_2O_3 (0001),[129] GaN/6H-SiC (0001),[130] AlGaAs/GaAs(001),[131] InGaAs/GaAs(001),[132,133] InGaAs/GaP (001),[130] InGaP/GaP (001),[130] ZnSe/GaAs(001),[134] ZnSe/Ge(001),[135,136] CdTe/InSb(001),[137] CdZnTe/GaAs(001),[138] CdTe/ZnTe/Si(112),[139] GaAs/Si(001),[140–144] wurtzite ZnS/Si(111),[145] diamond/Si(001),[146–148] and Si_3N_4/Si(111).[149]

Typically, if the substrate inclination is about an axis of symmetry, the tilt is about the same axis as the substrate inclination. Therefore, the surface normal, the low-index axis of the epitaxial layer, and the low-index axis of the substrate are coplanar. For this situation, the tilt is either away from (positive) or toward (negative) the surface normal. For pseudomorphic layers, the magnitude of the tilt increases with both the substrate inclination and the lattice mismatch. The tilt is positive (away from the surface normal) if $a_e > a_s$ but negative if $a_e < a_s$. In partially relaxed layers, the sign of the tilt is usually the opposite. However, the dependence of the tilt on the substrate inclination and mismatch is rather complex and incompletely understood at the present time.

5.7.1 Nagai Model

In the case of pseudomorphic growth, with no misfit dislocations at the interface, the tilt can be predicted by the Nagai model,[132] which can be understood with the aid of Figure 5.57. The vicinal substrate is assumed to comprise terraces of uniform length L separated by steps of height h. If the substrate inclination is Φ, then

$$\Phi = \tan^{-1}\left(\frac{h}{L}\right). \tag{5.124}$$

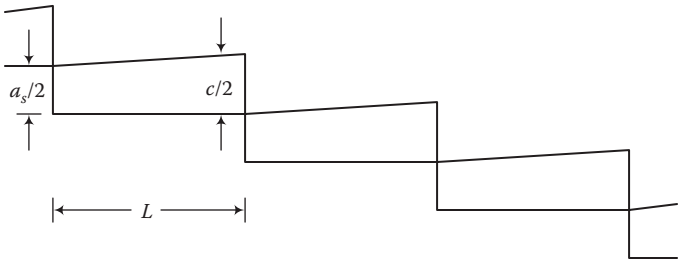


FIGURE 5.57
Nagai's model for tilting in a pseudomorphic heteroepitaxial layer deposited on a vicinal substrate. The vicinal substrate has uniform steps of height $a_s/2$ and separation L. For the case shown here, $c > a_s$, resulting in tilt away from the substrate normal (positive tilt). (Adapted from J. E. Ayers et al., *J. Cryst. Growth*, 113, 430 [1991]. With permission. Copyright 1991, Elsevier.)

This equation applies for single, double, or other step heights, as long as the step height is uniform across the wafer. For specificity, the steps were assumed to have a height of $a_s/2$ for the creation of Figure 5.57. Further, the epitaxial layer was assumed to be a cubic crystal, but tetragonally distorted (by the applied biaxial stress) with unit cell dimensions $a \times a \times c$. The substrate was assumed to be cubic and unstrained, with a lattice constant a_s. If coherency is maintained at the steps, so that the lattice constant of the epitaxial layer relaxes from a_s to c over the length of the terrace, then the epitaxial layer will be tilted with respect to the substrate by an amount $\Delta\Phi$ given by*

$$\Delta\Phi = \tan^{-1}\left(\frac{c - a_s}{a_s}\tan\Phi\right). \tag{5.125}$$

Although the Nagai model was developed for small offcut angles, Auvray et al.[150] reported that the general expression above is applicable to larger angles as well. This model predicts that the direction of tilt will be away from the surface normal (positive tilt) in the case of $c > a_s$ (or $a_e > a_s$), but toward the surface normal if $c < a_s$ (or $a_e < a_s$). The magnitude of tilt is predicted to increase with the substrate misorientation and lattice mismatch, as has been observed for pseudomorphic layers. Although this model was developed to explain tilting in zinc blende crystals, it should apply to hexagonal semiconductors as long as the correct biaxial relaxation constant is used to calculate the out-of-plane lattice constant.

In general, the introduction of dislocations at the interface will modify the tilt from the value predicted by the Nagai model. This will be true if the dislocations have Burgers vectors that are inclined to the interface. Here, the edge component of the Burgers vector that is normal to the interface can be considered a tilt component. (In contrast, the edge component of the Burgers vector that is in the plane of the interface is a misfit-relieving component.)

5.7.2 Olsen and Smith Model

Olsen and Smith[133] proposed a model to explain the tilting of a heteroepitaxial zinc blende semiconductor due to the introduction of misfit dislocations with Burgers vectors inclined to the growth interface. Suppose one type of dislocation is involved, with a tilt component (edge component perpendicular to the interface) b_1 and a misfit component (edge component parallel to the interface) b_2. Then, if the linear density of dislocations is just sufficient to relax the strain in the mismatched layer, the absolute value of the tilt will be approximately

$$|\Delta\Phi| \approx f\frac{b_1}{b_2}, \tag{5.126}$$

where f is the lattice mismatch. This expression is approximate because it does not consider the b_1 component necessary to relieve the lattice mismatch at the steps.

* The sign conventions used here differ from those sometimes used in the literature. In Equations 5.123 and 5.124, the substrate inclination is considered to always be positive; thus, the value of ϕ contains no information about the direction of this inclination. Further, the tilt of the epitaxial layer is considered positive if it adds to the substrate inclination but negative if it subtracts from it. Using these conventions, Equation 5.124 will correctly predict the sign of $\Delta\phi$.

There are two important limitations to the Olsen and Smith model. First, misfit dislocations exist in a two-dimensional array in the interface. Therefore, it is not possible to predict the direction of the tilt. Second, the Olsen and Smith model only predicts an upper bound for the tilt. This is because dislocations on different slip systems will have different b_1 components. If some are negative while some are positive, there will be partial cancellation, which will reduce the magnitude of the tilt. A more complete model for the crystallographic tilting of partially relaxed heteroepitaxial layers should take into consideration all of the active slip systems.

5.7.3 Ayers, Ghandhi, and Schowalter Model

Ayers, Ghandhi, and Schowalter[151] presented one such model for (001) heteroepitaxy of zinc blende semiconductors. Here, it was assumed that the relaxation was by 60° dislocations on {111}-type glide planes for layers greater than the critical layer thickness. Dislocation glide was modeled using the kinetic relaxation model of Matthews, Mader, and Light.[43] It was shown that the tilting of the substrate would create an asymmetry in the resolved shear stresses on the various slip systems. Because of this, preferential glide of dislocations on certain slip systems would lead to the crystallographic tilting of a heteroepitaxial layer on a vicinal substrate. This model is summarized below.

The eight active slip systems for (001) heteroepitaxy of zinc blende semiconductors are summarized in Table 5.2. Slip systems 1–4 involve α dislocations, while slip systems 5–8 involve β dislocations. In the case of an exact (001) substrate, the {111} glide planes all meet the interface at an angle of 54.7° along ⟨110⟩ directions. For each type of 60° dislocation, the Burgers vector contains a tilt component b_1 and a misfit component b_2. Due to the symmetry, the eight slip

TABLE 5.2

Eight Active Slip Systems for the (001) Heteroepitaxy of Zinc Blende Semiconductors

System	Line Vector l	Glide Plane	Burgers Vector b	Type of Dislocation
S_1	$[1\bar{1}0]$	(111)	$\frac{1}{2}a[10\bar{1}]$	α
S_2	$[1\bar{1}0]$	(111)	$\frac{1}{2}a[01\bar{1}]$	α
S_3	$[1\bar{1}0]$	$(\bar{1}\bar{1}1)$	$\frac{1}{2}a[101]$	α
S_4	$[1\bar{1}0]$	$(\bar{1}\bar{1}1)$	$\frac{1}{2}a[011]$	α
S_5	$[110]$	$(\bar{1}11)$	$\frac{1}{2}a[101]$	β
S_6	$[110]$	$(\bar{1}11)$	$\frac{1}{2}a[01\bar{1}]$	β
S_7	$[110]$	$(1\bar{1}1)$	$\frac{1}{2}a[10\bar{1}]$	β
S_8	$[110]$	$(1\bar{1}1)$	$\frac{1}{2}a[011]$	β

systems are all identically stressed and will contribute equal numbers of dislocations as the heteroepitaxial layer relaxes. Therefore, the tilt components of their Burgers vectors will cancel and there will zero net tilt of the epitaxial layer. In the case of a vicinal substrate, for which the [001] axis is inclined from the normal by an angle of Φ, this is no longer true.

The effect of various Burgers vectors components may be understood with the aid of Figure 5.58. The four dislocations shown all have pure edge character. The line vectors are each into the plane of the paper. Clockwise, Burgers circuits have been drawn in each case for the determination of the Burgers vector. It can be seen that the pure misfit dislocation of Figure 5.58a with its Burgers vector to the right will relieve mismatch strain in a layer with $a_e < a_s$ (tensile strain), whereas that of Figure 5.58b will relieve compressive strain. The tilt dislocation of Figure 5.58c with its Burgers vector "up" introduces clockwise tilt, but the dislocation of Figure 5.58d with its Burgers vector "down" causes counterclockwise tilt in the overlying crystal. (A screw component will neither relieve misfit nor introduce a macroscopic tilt in the epitaxial layer.)

The 60° dislocations in a heteroepitaxial zinc blende layer contain misfit, tilt, and screw components; however, we can still use the same principles outlined above to understand their behavior. In the case of a layer with tensile strain, the dislocations will be introduced with misfit components to the right. Then, with a counterclockwise substrate inclination, as shown in Figure 5.59a, dislocations with Burgers vector b_2 will be more stressed than those with Burgers vector b_1. The preferential introduction of b_2 dislocations will introduce positive tilt, which adds to the substrate inclination. Figure 5.59b shows the situation for

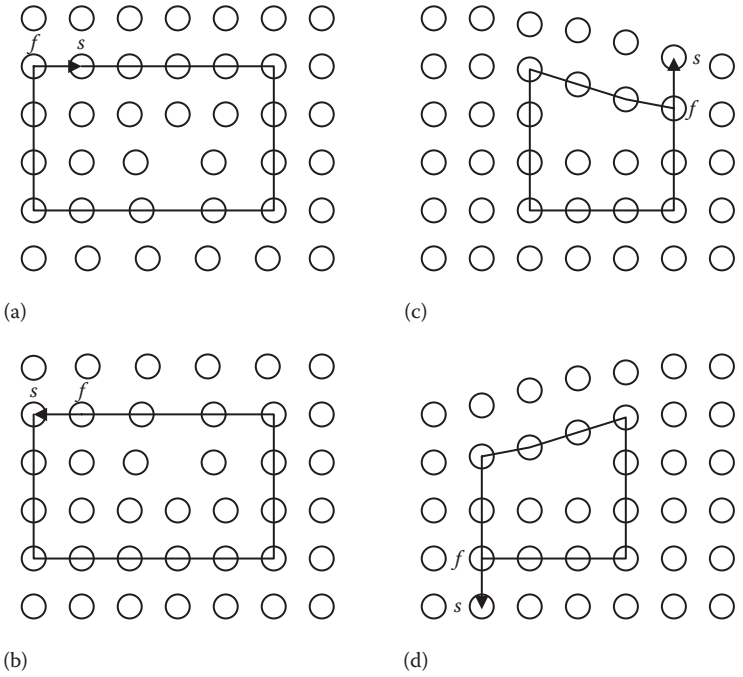


FIGURE 5.58
Pure misfit dislocations (a, b) and pure tilt dislocations (c, d) with edge character. In each case, the dislocation line vector is into the page, and a clockwise Burgers circuit is drawn from s to f. The Burgers vector is $\overline{fs}$. The dislocation shown will (a) relieve strain in an epitaxial layer with $a_e < a_s$ (tensile strain), (b) relieve strain in an epitaxial layer with $a_e > a_s$ (compressive strain), (c) introduce clockwise tilt, and (d) introduce counterclockwise tilt. (Adapted from J. E. Ayers et al., *J. Cryst. Growth*, 113, 430 [1991]. With permission. Copyright 1991, Elsevier.)

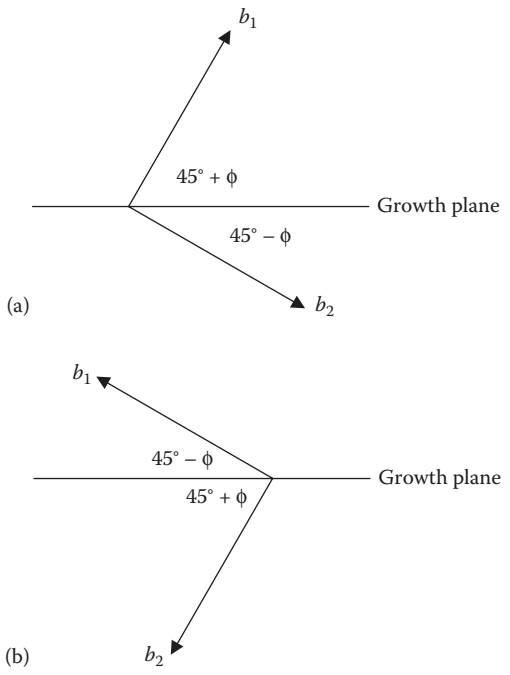


FIGURE 5.59
Burgers vectors of 60° dislocations in a heteroepitaxial zinc blende semiconductor on a vicinal (001) substrate. (a) Tensile mismatch, with a counterclockwise substrate inclination. (b) Compressive mismatch, with a counterclockwise substrate inclination. (Adapted from J. E. Ayers et al., *J. Cryst. Growth*, 113, 430 [1991]. With permission. Copyright 1991, Elsevier.)

a counterclockwise tilt but compressive mismatch. Here, the strain must be relieved by dislocations with their misfit components to the left. The preferential introduction of b_1 dislocations, which are more stressed in this case, will introduce a negative tilt, which subtracts from the substrate inclination.

The quantitative determination of the tilt in the heteroepitaxial layer requires (1) the determination of the densities of dislocations on the eight slip systems and (2) the summing of their contributions to the epitaxial layer tilt. This was done for two limiting cases. In the case of type I relaxation, it was assumed that all eight slip systems would become active, most of the relaxation would occur with $h \gg h_c$, and the more stressed systems would contribute more misfit dislocations. In the case of type II relaxation, it was assumed that the relaxation would be affected only by the most stressed slip systems (MSSSs) for the two ⟨110⟩ directions. In other words, the least stressed slip systems (LSSSs) are excluded as a consequence of relaxation by the others. This could be caused by differences in critical thickness, glide, multiplication, or nucleation of the dislocations on the different slip systems. In the limiting case of type II relaxation, the tilt would be equal to that predicted by the Olsen and Smith model, but now the direction of tilt may be predicted.

For example, the imbalance of the dislocation populations could be the result of differences in the critical layer thicknesses (which arise from the substrate inclination) for the dislocations in the different slip systems.[152] The MSSSs, which have a lower critical layer thickness, will initiate relaxation by glide before the LSSSs. After the MSSSs become active, they can continually reduce the strain in the growing layer, thus keeping the LSSSs inactive. In such a situation, the LSSSs may be completely excluded from the relaxation process.

Tsao and Dodson[153] have shown that a slip system will become active (introduce misfit dislocations to relax strain) only when its effective stress τ_{eff} becomes positive, where

$$\tau_{eff} = \left| \frac{2G\varepsilon(1+\nu)}{1-\nu} \right| - \frac{Gb}{2\pi h}\left[\frac{1-\nu\cos^2\alpha}{1-\nu}\right]\left[\ln\left(\frac{4h}{b}\right)\right], \tag{5.127}$$

and
where:

- λ is the angle between the slip direction and that direction in the plane of the interface that is perpendicular to the intersection of the glide plane and the interface
- G is the shear modulus
- ε is the average strain in the epitaxial layer
  is the Poisson ratio
- α is the angle describing the dislocation character (60°)
- h is the layer thickness
- b is the length of the Burgers vector

In the case of a vicinal substrate, the different slip systems will have different values of τ_{eff} due to the different values of λ. Type II relaxation is affected entirely by the MSSSs in one of two scenarios. In the first, the relaxation takes place near equilibrium, so that the MSSSs maintain $\tau_{eff}=0$. Then τ_{eff} is negative for the LSSSs, and they will not participate in relaxation. In the second (more likely) situation, the relaxation is limited by kinetics and does not take place near equilibrium. The LSSSs will experience less negative, or perhaps even positive, values of τ_{eff}. Nonetheless, all slip systems are subject to essentially the same kinetic limitations. This means the MSSSs can still maintain values of the excess stress that are much greater than those for the LSSSs, and the MSSSs can relieve the strain while essentially excluding the LSSSs.

For the case of Type I relaxation, the dislocation populations were estimated using the Matthews, Mader, and Light kinetic model for lattice relaxation described in Section 5.6.5. In the Type I case, all slip systems are active, and so the calculation of the tilt requires the determination of their relative contributions. In contrast to the Type II case, it is therefore necessary to assume a limiting mechanism for lattice relaxation by the individual slip systems. For this model, it is assumed that lattice relaxation is limited by the glide of dislocations. Dislocations on slip systems S_1 through S_4 relieve strain in the $[1\bar{1}0]$ direction, while S_5 through S_8 are associated with strain relief in the [110] direction. If δ_i is the strain relaxation by dislocations on the ith slip system, then the strains in the two <110> directions are

$$\varepsilon[110] = -f + (\delta_1 + \delta_2 + \delta_3 + \delta_4) \tag{5.128}$$

and

$$\varepsilon\left[1\bar{1}0\right] = -f + (\delta_5 + \delta_6 + \delta_7 + \delta_8). \tag{5.129}$$

The time rate of change of the lattice relaxation by the ith slip system is given by

$$\frac{d\delta_i}{dt} \approx \frac{2Gb^2(1+\nu)\varepsilon\rho_i b_{i2}\cos^2\lambda_i\cos\phi_i}{k_B T(1-\nu)\exp(U/k_B T)}, \tag{5.130}$$

where:

- G is the shear modulus
- b is the length of the Burgers vector
- ν is the Poisson ratio
- ε is the strain in the appropriate <110> direction
- ρ_i is the linear density of misfit dislocations
- b_{i2} is the misfit component of the Burgers vector
- λ_i is the angle between the slip direction and that direction in the plane of the film that is perpendicular to the intersection of the glide plane and the plane of the film
- ϕ_i is the angle between the film surface and the normal to the slip plane
- U is the activation energy for dislocation glide

The geometric factors for the eight slip systems can be found as follows. If the vicinal (001) substrate is inclined α degrees toward the [100] and β degrees toward the [010], then the substrate unit normal is

$$\hat{n} = \left[\sin\alpha, \sin\beta, \left(1 - \sin^2\alpha - \sin^2\beta\right)^{1/2}\right]. \tag{5.131}$$

If $\hat{b}_i$ is the unit vector in the direction of the Burgers vector and $\hat{g}_i$ is the unit normal to the glide plane, then

$$\cos\lambda_i = \frac{b_i \cdot \left[n \times \left(n \times g_i\right)\right]}{\left|b_i \cdot \left[n \times \left(n \times g_i\right)\right]\right|}. \tag{5.132}$$

If the substrate inclination is small, then the direction cosines for the eight slip systems are given approximately by

$$\cos\lambda_1 \approx \frac{1+\beta}{2-\alpha-\beta},$$

$$\cos\lambda_2 \approx \frac{1+\alpha}{2-\alpha-\beta},$$

$$\cos\lambda_3 \approx \frac{1-\beta}{2+\alpha+\beta},$$

$$\cos\lambda_4 \approx \frac{1-\alpha}{2+\alpha+\beta},$$

$$\cos\lambda_5 \approx \frac{1+\beta}{2+\alpha-\beta},$$

$$\cos\lambda_6 \approx \frac{1-\alpha}{2+\alpha-\beta},$$

$$\cos\lambda_7 \approx \frac{1-\beta}{2-\alpha+\beta},$$

and

$$\cos\lambda_8 \approx \frac{1+\alpha}{2-\alpha+\beta}. \tag{5.133}$$

Similarly,

$$\phi_i = \sin^{-1}\left\{\left|g_i \cdot n\right|\right\},$$

so that

$$\phi_1 = \phi_2 = \sin^{-1}\left\{\frac{\left[\sin\alpha+\sin\beta+\left(1-\sin^2\alpha-\sin^2\beta\right)^{1/2}\right]}{\sqrt{3}}\right\},$$

$$\phi_3 = \phi_4 = \sin^{-1}\left\{\frac{\left[-\sin\alpha-\sin\beta+\left(1-\sin^2\alpha-\sin^2\beta\right)^{1/2}\right]}{\sqrt{3}}\right\},$$

$$\phi_5 = \phi_6 = \sin^{-1}\left\{\frac{\left[-\sin\alpha+\sin\beta+\left(1-\sin^2\alpha-\sin^2\beta\right)^{1/2}\right]}{\sqrt{3}}\right\},$$

$$\phi_7 = \phi_8 = \sin^{-1}\left\{\frac{\left[\sin\alpha-\sin\beta+\left(1-\sin^2\alpha-\sin^2\beta\right)^{1/2}\right]}{\sqrt{3}}\right\}. \tag{5.134}$$

The magnitude and direction of the tilt can be calculated if the misfit strain relaxed by each slip system (the δ_i) is known. A slip system with line vector l_1 will produce tilt about the l_1 axis but relieve strain in the l_2 direction. Also, a slip system with line vector l_2 will produce tilt about the l_2 axis but relieve strain in the l_1 direction. The strain relaxed by each slip system was estimated as follows.

The values of ρ_i are assumed to be approximately equal, so the relaxation rates for two different slip systems will be in the ratio

$$\frac{d\delta_i/dt}{d\delta_j/dt} \approx \frac{\cos^2\lambda_i \cos\phi_i}{\cos^2\lambda_j \cos\phi_j}. \tag{5.135}$$

Then, if nearly complete relaxation has occurred, the lattice relaxation by the *i*th slip system can be found from

$$\delta_i \approx \frac{f\cos^2\lambda_i\cos\phi_i}{\sum_{j=1,4}\cos^2\lambda_j\cos\phi_j}, \text{ for } i = 1, 2, 3, 4, \tag{5.136}$$

and

$$\delta_i \approx \frac{f\cos^2\lambda_i\cos\phi_i}{\sum_{j=5,8}\cos^2\lambda_j\cos\phi_j}, \text{ for } i = 5, 6, 7, 8. \tag{5.137}$$

The resulting tilt can be calculated by combining the contributions of the eight slip systems, as follows. If γ and η are the tilts about the $[\bar{1}10]$ and [110] axes, respectively, then in the case of complete relaxation,

$$\gamma = \tan^{-1}\left\{ f\tan\left[\sqrt{2}(\alpha+\beta)\right] - \sum_{i=1,4}\frac{\delta_i b_{i1}}{b_{i2}}\right\} \tag{5.138}$$

and

$$\eta = \tan^{-1}\left\{ f\tan\left[\sqrt{2}(\alpha-\beta)\right] - \sum_{i=5,8}\frac{\delta_i b_{i1}}{b_{i2}}\right\}. \tag{5.139}$$

The "intrinsic" tilt due to the surface steps has been included here, and it will cancel the dislocation contributions at least in part. Finally, the overall tilt can be found from

$$\Delta\Phi = \cos\left[\left(1+\tan^2\gamma+\tan^2\eta\right)^{-1/2}\right]. \tag{5.140}$$

As a result, the tilt is predicted to be (approximately) proportional to the substrate inclination and the lattice mismatch, as has been observed. The predictions of this model for the cases of Type I and Type II relaxation are shown in Figure 5.60, for the case of 2° substrate inclination, along with experimental data from several epitaxial systems.

This model successfully predicts the direction of the tilt for tensile and compressive layers, both pseudomorphic ($\delta_i = 0$) and relaxed. It also introduced a framework for the quantitative prediction of the absolute tilt. As a result, the tilt was predicted for the two limiting cases of Type I relaxation (all eight slip systems active) and Type II relaxation (only the most stressed systems active.) The key limitation of the model is that it did not account for either dislocation nucleation or multiplication in the Type I case, although both are known to be important in determining the dynamics of lattice relaxation. However, the effect of these phenomena, either acting alone or in conjunction, will be to exaggerate the imbalance in lattice relaxation among the eight slip systems. It will therefore produce a tilt

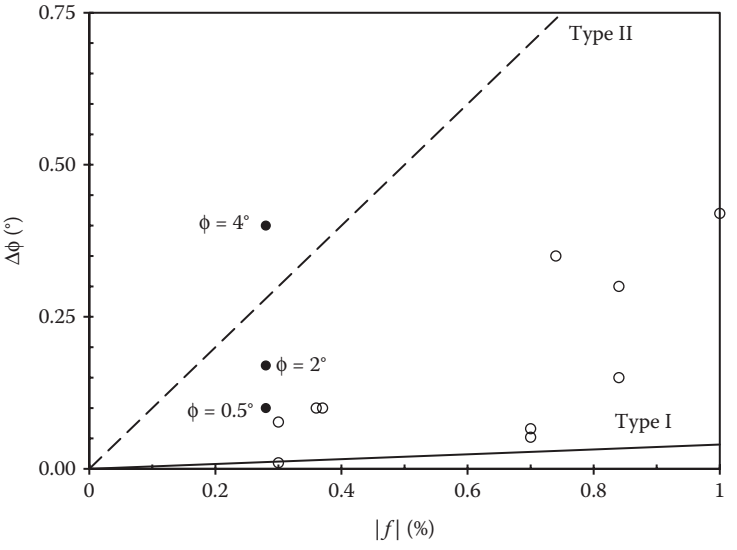


FIGURE 5.60
Tilt ΔΦ vs. the absolute value of the lattice mismatch f for (001) heteroepitaxy of zinc blende semiconductors. The two lines were calculated for the Type I and Type II limiting cases, as indicated, for a substrate inclination of 2°. The experimental data shown by open circles were for substrates having a 2° inclination. For the filled circles, the substrate inclination was as indicated. (Adapted from J. E. Ayers et al., *J. Cryst. Growth*, 113, 430 [1991]. With permission. Copyright 1991, Elsevier.)

that is greater than the Type I case, and closer to the Type II limit. And, as pointed out previously, the Type II case is mechanism independent, as long as the MSSSs relax the strain at the exclusion of the others, and no new slip systems become active. Therefore, the Type I and Type II cases can still be considered to give the minimum and maximum tilt that should be expected for (001) heteroepitaxy of zinc blende semiconductors. Nonetheless, a far more detailed model is needed in order to make accurate quantitative predictions of the tilt or perhaps to use the observed tilts to extract information regarding the dislocation dynamics.

5.7.4 Riesz Model

Riesz[154] extended the model of Ayers, Ghandhi, and Schowalter to include dislocation multiplication, using the approach of Dodson and Tsao,[101] for the (001) heteroepitaxy of zinc blende semiconductors. Here, it was assumed that two types of slip systems are active: the MSSSs, called set A, and the LSSSs, called set B.* For each set, the dislocation multiplication was assumed to be described by

$$\frac{\partial \rho}{\partial t} = C\left(\frac{\tau_{\text{eff}}}{G}\right)^2 \left[\rho(t) + \rho_0\right], \tag{5.141}$$

* Riesz considered the case of substrate inclination toward a <100> direction (α = 0 or β = 0), for which there are only two distinct values of λ among the eight slip systems. In the general case (both α and β nonzero), there will be four distinct values of λ among the eight slip systems. This will complicate the model considerably, but is not expected to change the qualitative results.

where:

τ_{eff} is the effective stress
G is the shear modulus
ρ is the linear density of misfit dislocations
C is a thermally activated factor given by

$$C = C_0 \exp(-U/k_B T), \tag{5.142}$$

where U is the activation energy for dislocation glide. The dislocation multiplication processes for the A and B dislocations are assumed to be independent, because dislocation multiplication sources usually emit dislocations with the same Burgers vector.

The strain relaxation in either <110> direction is due to the combined effect of dislocations from sets A and B; hence,

$$\delta = b(\rho_A \sin\lambda_A \sin\phi_A + \rho_B \sin\lambda_B \sin\phi_B). \tag{5.143}$$

The resulting tilt $\Delta\Phi$ of the epitaxial layer may be calculated from

$$\Delta\Phi = \Phi\left[f\left(\frac{1+\nu}{1-\nu}\right) - \delta\left(\frac{2\nu}{1-\nu}\right)\right] - b\sin\lambda\sin\phi(\rho_A - \rho_B), \tag{5.144}$$

where Φ is the inclination of the vicinal substrate. Here, the first term is the "Nagai contribution" due to steps at the interface and the second term is due to the asymmetric lattice relaxation by the dislocations from sets A and B.

Using this model, the epitaxial layer tilt $\Delta\Phi$ was calculated as a function of the growth temperature, with the initial dislocation density ρ_0 as a parameter. For these calculations, it was assumed that $C_0 = 5\times10^{11}\,s^{-1}$ and $U = 1.7$ eV. The results of these calculations are shown in Figure 5.61 for substrate miscut angles of 0.2°, 2°, and 4°. Also shown in the figure are the Type I and Type II limits predicted by the Ayers, Schowalter, and Ghandhi model.

The tilt behavior is predicted to be intermediate between the Type I and Type II limits, as expected. In all cases, the predicted tilt increases with the growth temperature. For substrate inclinations of 2° or 4°, Type II behavior is approached for high growth temperatures. This shows that, if the MSSSs have sufficiently fast slide and multiplication processes, they may largely exclude the other slip systems.

Type I behavior was approximated over much of the temperature range if the substrate inclination was small. This is to be expected; in the limit of zero substrate inclination, the lattice relaxation is symmetric so all slip systems will participate. With larger substrate inclinations, however, Type I behavior was predicted only if the initial dislocation densities were excessive.

The tendency toward Type I or Type II behavior therefore appears to be controlled by the starting dislocation density, as well as the substrate inclination and growth temperature. This is further illustrated in Figure 5.62. Here, the epitaxial layer tilt is plotted as a function of the substrate inclination, with the initial defect density as a parameter. The behavior is approximately Type I over the range of miscut angles only if the initial dislocation density is very high. Otherwise, there is a tendency toward Type II behavior as the substrate inclination, and therefore the asymmetry between the slip systems, is increased.

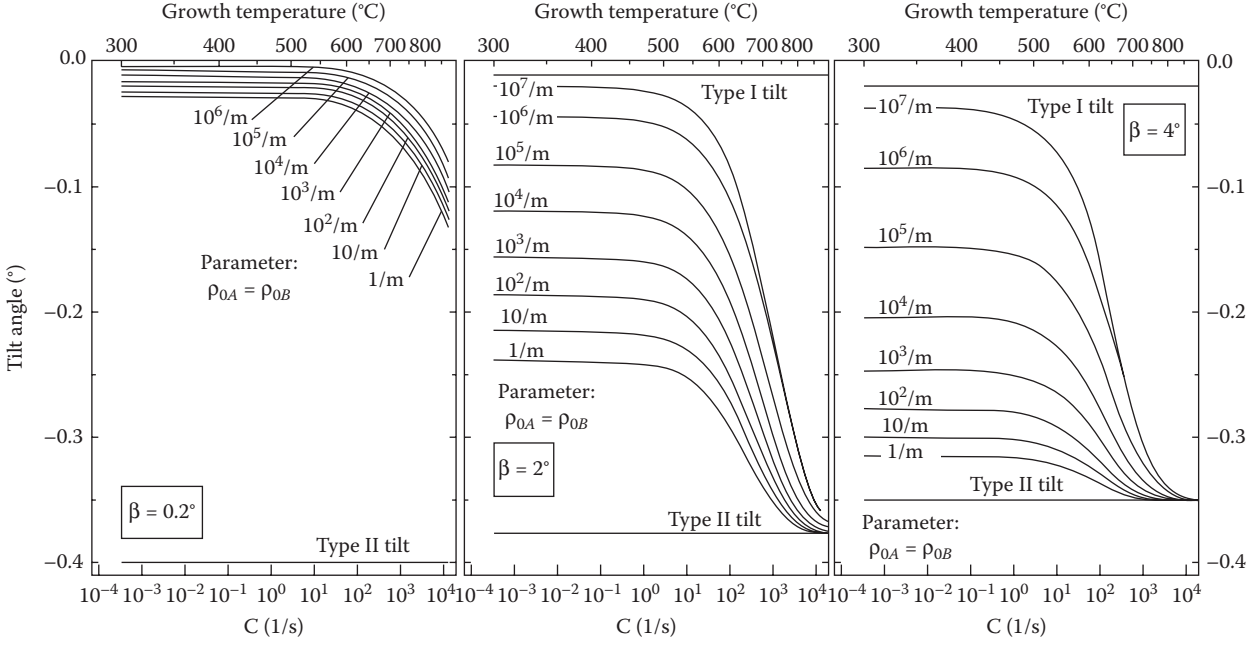


FIGURE 5.61
Predicted tilt ΔΦ vs. growth temperature, with the initial dislocation density ρ_0 as a parameter. It was assumed that $C_0 = 5\times10^{11}\text{s}^{-1}$ and $E_A = 1.7\text{ eV}$. The substrate inclination was assumed to be 0.2°, 2°, and 4° for the first, second, and third graphs, respectively. Also shown are the Type I and Type II limits predicted by the Ayers, Schowalter, and Ghandhi model. (Reprinted from F. Riesz, *J. Appl. Phys.*, 79, 4111 [1996]. With permission. Copyright 1996, American Institute of Physics.)

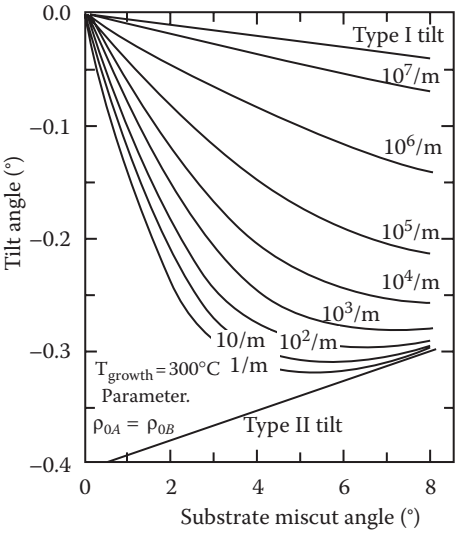


FIGURE 5.62
Predicted tilt $\Delta\Phi$ as a function of the substrate inclination, with the initial dislocation density as a parameter. It was assumed that $C_0 = 5\times10^{11}\text{s}^{-1}$, $E_A = 1.7$ eV, and $T = 300°C$. (Reprinted from F. Riesz, *J. Appl. Phys.*, 79, 4111 [1996]. With permission. Copyright 1996, American Institute of Physics.)

The characteristics of Figures 5.61 and 5.62 were calculated with the assumption of complete lattice relaxation. In partially relaxed layers, the tilt varies monotonically as the extent of lattice relaxation increases. This is shown in Figure 5.63, for the case of a substrate inclination $\Phi = 4°$, with $\rho_{0A} = \rho_{0B} = 10^3\text{m}^{-1}$. In the pseudomorphic case ($\delta = 0$), a small positive tilt is observed due to the interfacial steps (Nagai contribution). For the partially relaxed layers, the tilt varies monotonically with the extent of the relaxation.

In summary, the tilts observed in zinc blende and diamond heteroepitaxial layers are caused by asymmetric relaxation by the different slip systems. The asymmetric relaxation arises from differences in both dislocation glide and multiplication with an inclined substrate. The overall behavior is intermediate between the Type I and Type II limits. (Type I relaxation involves relaxation by all of the slip systems, with asymmetries introduced by weak geometric factors. Type II relaxation involves relaxation by only the MSSSs, at the exclusion of the others.) Type I behavior is favored only for very high initial dislocation densities or small substrate inclinations. Type II behavior is approached as the substrate inclination or temperature is increased.

5.7.5 Vicinal Epitaxy of III-Nitride Semiconductors

The III-nitrides such as AlN and GaN have typically been grown on vicinal SiC or sapphire substrates, and it has been found that vicinal surface epitaxy (VSE) results in heteroepitaxial layers of improved crystal quality with either type of substrate.[155–159]

In significantly relaxed (metamorphic) nitride layers grown on vicinal surfaces, the tilts are as predicted by the Nagai model, even though this model was intended to apply in the pseudomorphic case. This has been found to be the case for AlN/6H-SiC (0001) and also for GaN/Al_2O_3 (0001) with small offcut angles. The same is true for heteroepitaxial GaN on AlN, when the AlN is a buffer layer grown on vicinal SiC (0001).

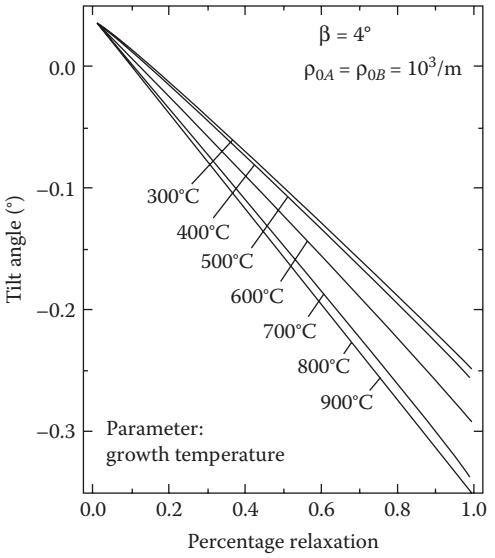


FIGURE 5.63
Predicted tilt ΔΦ as a function of the extent of lattice relaxation. On the abscissa, zero represents the pseudomorphic case and 1.0 represents complete lattice relaxation. It was assumed that $C_0 = 5\times10^{11}\,s^{-1}$, $E_A = 1.7$ eV, and $\rho_{0A} = \rho_{0B} = 10^3\,m^{-1}$. (Reprinted from F. Riesz, *J. Appl. Phys.*, 79, 4111 [1996]. With permission. Copyright 1996, American Institute of Physics.)

For diamond and zinc blende semiconductors, the Nagai model only applies to the pseudomorphic case. This is because the 60° misfit dislocations on $a/2\langle011\rangle\{111\}$ slip systems have Burgers vectors with tilt components, which in general do not cancel. However, for the (0001) heteroepitaxy of III-nitrides, the misfit dislocations have in-plane Burgers vectors. Therefore, they do not affect the crystallographic tilting of the heteroepitaxial layer, and Nagai's model should apply to pseudomorphic, partially relaxed, or even fully relaxed layers.

Huang et al.[130] studied the crystallographic tilting in MOVPE-grown GaN/AlN/6H-SiC (0001) heterostructures using TEM and x-ray diffraction. They compared layers grown on two types of substrates: exact (0001) and $(0001)\,3.5° \rightarrow [11\bar{2}0]$. For the case of the vicinal substrate, they found that the AlN was tilted with respect to the substrate by 142 in., consistent with the Nagai model and the measured change in the out-of-plane lattice constant, $\Delta c/c = -1.05\%$. Also, the tilting of the GaN overlayer with respect to the AlN buffer was –370 in., also consistent with the Nagai model and the measured change in the out-of-plane lattice constant for that interface, which was $\Delta c/c = 3.94\%$.

The agreement between the Nagai model and the experimental tilts indicates that the misfit dislocations at both the GaN/AlN (0001) and AlN/6H-SiC (0001) interfaces have in-plane Burgers vectors (*a*-type dislocations). This was confirmed for the AlN/SiC interface using cross-sectional HRTEM images. The $[1\bar{1}00]$ projection HRTEM image was filtered using a masked fast Fourier transform. The $11\bar{2}0$-filtered image revealed misfit dislocations with extra half planes in the SiC substrate. However, the 0004-filtered image showed no vertical displacement around the misfit dislocations, demonstrating that their Burgers vectors were indeed within the plane of the interface.

GaN on sapphire (0001) shows a similar behavior, for small values of substrate inclination. Huang et al. studied the crystallographic tilting in MOVPE-grown GaN/Al_2O_3 (0001).

The substrate inclination and the epitaxial layer tilt, with respect to the substrate, were determined using back-reflection synchrotron Laue x-ray diffraction patterns. For the case of a (0001) $1.85° \rightarrow [11\bar{2}0]$ substrate, the direction and magnitude of the tilt were in agreement with the Nagai model. This result is taken to mean that the misfit dislocations have zero tilt components (they are *a*-type dislocations, with Burgers vectors in the plane of the interface). For larger substrate inclinations, the tilting behavior was quite different. However, this is believed to be affected by step heights greater than two atomic layers on these substrates. (This is described in greater detail in Section 5.7.6.)

The tilting behavior of the hexagonal III-nitrides on 6H-SiC and sapphire substrates indicates that the misfit dislocations have zero tilt components. In other words, these MDs have in-plane Burgers vectors. Unless the MDs come about by the reaction of dislocations having out-of-plane Burgers vectors, they must be introduced by basal-plane slip (slip on the (0001) plane). The former possibility can probably be ruled out; otherwise, some unreacted dislocations with out-of-plane Burgers vectors should have been observed. If basal-plane slip is the dominant means for the introduction of MDs in III-V nitrides on *c*-face substrates, this is in sharp contrast with the diamond and zinc blende semiconductors for which MDs glide to the interface on {111}-type planes.

5.7.6 Vicinal Heteroepitaxy with a Change in Stacking Sequence

An interesting feature of AlN/6H-SiC (0001) heteroepitaxy is that these two crystals have different stacking sequences in the growth direction. In the [0001] direction, the stacking sequence for the 6H-SiC substrate is ABCACBA, but for the wurtzite AlN, it is ABA. (The wurtzite structure can be considered to have a 2H stacking sequence.) If a vicinal substrate is used, defects must be introduced at the interfacial steps to accommodate the change in stacking sequence.

Huang et al.[130] studied the misfit dislocation structure in MOVPE-grown GaN/AlN/6H-SiC (0001) heterostructures using TEM. The vicinal 6H-SiC substrates were (0001) and $(0001)\ 3.5° \rightarrow [11\bar{2}0]$. From the analysis of TEM results, they concluded that most of the misfit dislocations were 60° Shockley partial dislocations, which they called "geometrical partial misfit dislocations" (GPMDs). Because the GPMDs were generally on different terraces, they appeared to be unpaired partial dislocations, rather than the paired partials that would be expected if they arose from the dissociation of perfect dislocations on a basal glide plane.

Huang et al. showed that the high density of unpaired partial dislocations in their samples could serve to accommodate the difference in stacking sequence between the AlN epitaxial layer and its 6H-SiC substrate. The structural model they proposed is shown in Figure 5.64. Figure 5.64a depicts a side view of the interface showing the steps on the (0001) surface. (The plane of the paper is the $(1\bar{1}00)$ face.) Above the terraces, the AlN layer may take on the ABA or the ACA stacking sequence, either of which results in the wurtzite structure. A transition between these two stacking sequences at a step can be accommodated by the introduction of a 60° Shockley partial dislocation, as shown in Figure 5.64b. This transition is gradual and preserves the hexagonal structure of the AlN, albeit with some distortion. Moreover, the gradual change in the stacking sequence, as shown in Figure 5.64c, does not create vertical boundaries in the epitaxial layer.

These results for AlN/6H-SiC (0001) have interesting implications for vicinal substrate epitaxy (VSE) of III-nitrides. If the geometric partial dislocations are introduced to accommodate the difference in stacking sequences between the epitaxial layer and the substrate, then their introduction is controlled in part by the substrate inclination. This offcut angle

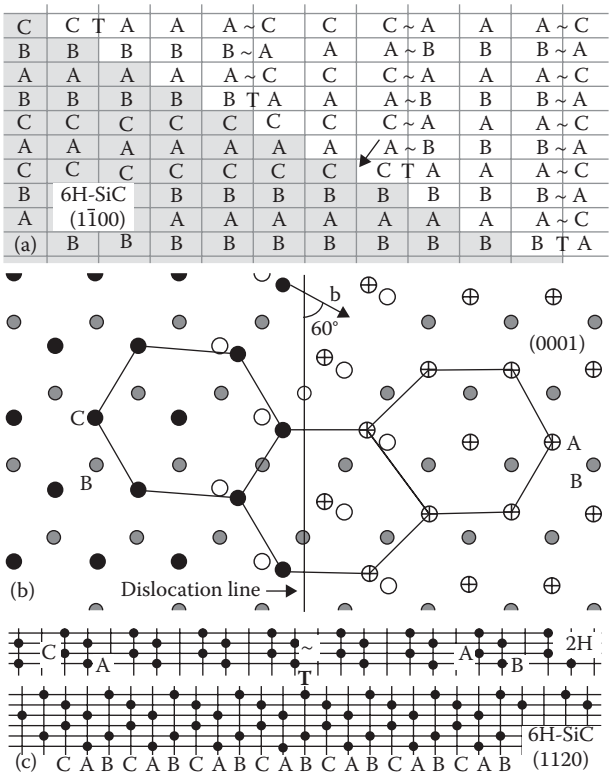


FIGURE 5.64
Possible arrangement for the accommodation of the stacking sequence difference at the AlN/6H-SiC (0001) interface by geometric partial misfit dislocations (GPMDs). (a) Partial dislocations are introduced at the steps. (b) The gradual transition preserves the hexagonal structure of the cells, with some distortion. (c) The gradual transition does not result in vertical boundaries within the epitaxial AlN. (Reprinted from X. R. Huang et al., *Phys. Rev. Lett.*, 95, 86101 [2005]. With permission. Copyright 2005, American Physical Society.)

might be used to effect the lattice relaxation, the introduction of dislocations, and the threading dislocation density in the heteroepitaxial material.

5.7.7 Vicinal Heteroepitaxy with Multilayer Steps

Up to now, we have only considered vicinal heteroepitaxy with monatomic or bilayer steps between the substrate terraces. However, step bunching can occur on some substrates with large miscut angles. For example, vicinal Al_2O_3 (0001) (*c*-face sapphire) substrates annealed at temperatures above 1200°C exhibit steps of *n*-bilayer height, with $1 < n < 6$.[160]

Huang et al.[129] investigated the tilting of GaN/Al_2O_3 (0001). They found that the tilt was in agreement with the Nagai model only for small substrate inclinations. For offcuts of 6.29° or 10.6° toward the [1$\bar{1}$00], the measured tilts were very different from those predicted by the Nagai model, and for the case of 10.6° inclination, the direction of the tilt was opposite that predicted.

Huang et al. explained their measured results using the schematics of Figure 5.65. Assuming that the misfit dislocations have in-plane Burgers vectors, the tilt of the epitaxial layer should be the same as predicted by the Nagai model for $n = 1$ or $n = 2$, as shown in Figure 5.65a and b, respectively:

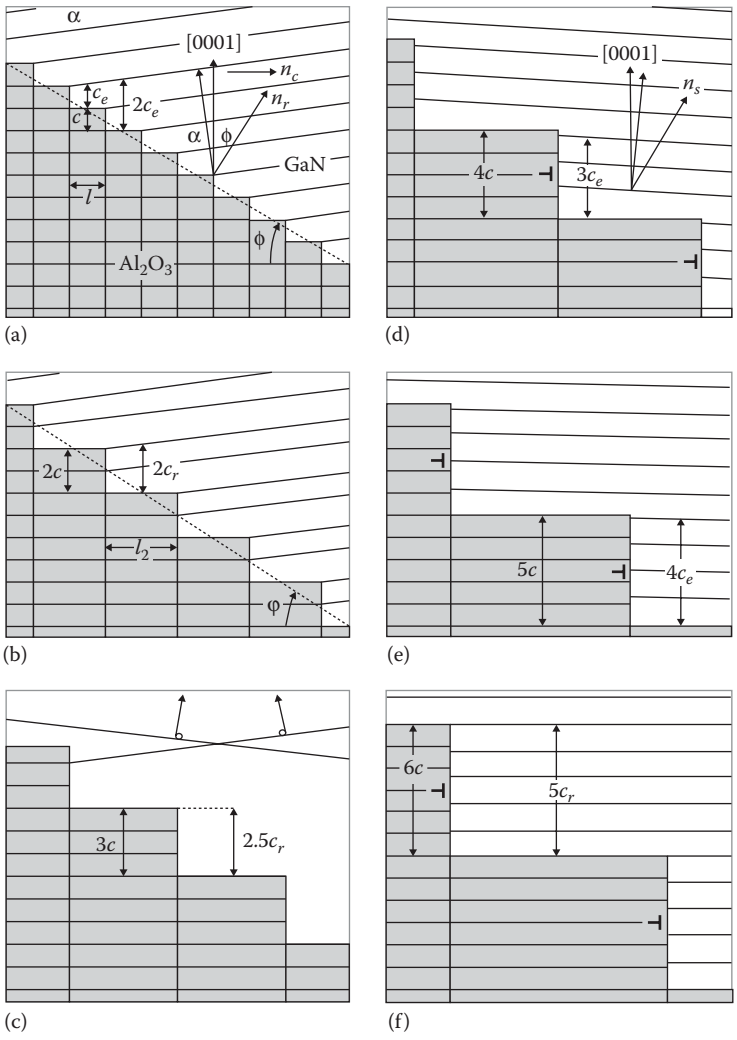


FIGURE 5.65
Schematic diagrams of GaN grown heteroepitaxially on vicinal sapphire (0001) having various step heights of nc, where c is the lattice constant of the substrate. (a) $n = 1$ (step height = c), (b) $n = 2$, (c) $n = 3$, (d) $n = 4$, (e) $n = 5$, (f) $n = 6$. c is the substrate lattice constant and c_e is the epitaxial layer lattice constant. n_s is the substrate surface normal and n_c is the offcut direction. (Reprinted from X. R. Huang et al., *Appl. Phys. Lett.*, 86, 211916 [2005]. With permission. Copyright 2005, American Institute of Physics.)

$$\Delta\Phi = \tan^{-1}\left(\frac{\Delta c}{c}\tan\Phi\right). \tag{5.145}$$

For a three-bilayer step as shown in Figure 5.65c, however, the local tilt will have a magnitude given by $|\Delta\Phi| = \tan^{-1}\{[(3c - 2c_e)/3c]\tan\Phi\}$. However, since this tilt can take on either sign with equal probability, the average tilt is expected to be zero. For the four-bilayer step of Figure 5.65d, the tilt would be

$$\Delta\Phi = -\tan^{-1}\left[\frac{(4c - 3c_e)}{4c}\tan\Phi\right]. \tag{5.146}$$

This tilt will have a sign opposite of that predicted by the Nagai model for four-bilayer steps. For the five-bilayer step,

$$\Delta\Phi = -\tan^{-1}\left[\frac{(5c - 4c_e)}{5c}\tan\Phi\right], \tag{5.147}$$

and for the six-bilayer step,

$$\Delta\Phi = -\tan^{-1}\left[\frac{(6c - 5c_e)}{6c}\tan\Phi\right]. \tag{5.148}$$

In this case, the approximate matching between $6c$ and $5c_e$ results in negligible tilt.

Thus, the Nagai model may be used for pseudomorphic layers or relaxed layers for which the misfit dislocations have in-plane Burgers vectors. But if the steps on the vicinal substrate are higher than two bilayers, it is necessary to consider the matching between different integral multiples of the substrate and epitaxial layer lattice constants.

5.8 Dislocation Coalescence, Annihilation, and Removal in Relaxed Heteroepitaxial Layers

In most thick, (nearly) relaxed heteroepitaxial layers, it is found that (1) the threading dislocation density greatly exceeds that of the substrate, and (2) this dislocation density (measured at the surface or averaged over the thickness) decreases approximately with the inverse of the thickness, as noted by Sheldon et al.[161] for a number of heteroepitaxial material systems. Figure 5.66 shows the observed threading dislocation densities as a function of layer thickness, for several heteroepitaxial systems.

To better understand this behavior, Ayers et al.[162] measured the threading dislocation densities in GaAs/Si(001) heterostructures grown by MOVPE, both as grown and after postgrowth annealing treatments. Uncracked layers with thicknesses up to about 4 μm were studied. For as-grown samples, the dislocation density was found to be inversely proportional to the layer thickness, as expected. The dislocation density could be reduced by postgrowth annealing. However, for all annealing temperatures investigated, the dislocation density saturated at a minimum value and could not be further reduced by additional annealing time. It is significant that *the minimum value of the dislocation density, which presumably represented some stable configuration of dislocations, was found to be inversely proportional to the thickness, but with a smaller constant of proportionality than for the as-grown samples*.

Several models have been proposed to explain these experimental results, but all of them involve dislocation–dislocation reactions that can reduce the threading dislocation density with increasing thickness. During the early stages of relaxation, new dislocations are created by heterogeneous nucleation and multiplication processes. But once most of the lattice

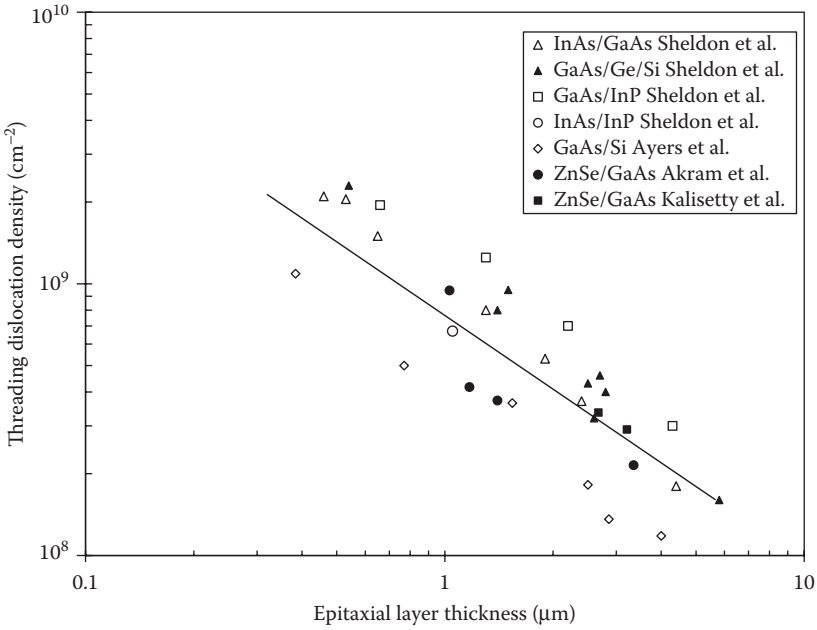


FIGURE 5.66
Threading dislocation density vs. epitaxial layer thickness, for several mismatched heteroepitaxial material systems. The data are from Sheldon et al.,[161] Ayers et al.,[162] Akram et al.,[163] and Kalisetty et al.,[164] as indicated in the legend.

mismatch has been relieved by misfit dislocations, the threading dislocations (which are nonequilibrium defects) can react with other threading dislocations, leading to coalescence or annihilation. In some cases, dislocations may glide to the edge of the sample and be removed in that way. However, this is only expected to be important in small (patterned) regions of heteroepitaxial material. Therefore, in a planar (unpatterned) layer, coalescence and annihilation are the important processes for dislocation removal. Here, coalescence refers to a reaction between two threading dislocations having different Burgers vectors; the end product is a single threading dislocation, and so one threading dislocation is removed. Annihilation refers to the reaction of two dislocations having antiparallel Burgers vectors, which leads to the removal of both. These processes, involving thermally activated glide of dislocations, will only occur during the growth itself or subsequent thermal processing.

A semiempirical model for dislocation coalescence and annihilation was developed by Tachikawa and Yamaguchi. The equation governing the reduction of the dislocation density D with the thickness h was assumed to include first-order and second-order dislocation interactions, so that

$$\frac{dD}{dh} = -C_1 D - C_2 D^2 , \tag{5.149}$$

where C_1 and C_2 are constants corresponding to the probabilities of each reaction. A coalescence reaction occurs when two dislocation approach within a reaction distance d_0 and therefore $C_2 = \rho d_0$ where ρ is the effective propagation constant due to the dislocation geometry and the growth direction. Annihilation may occur from the simultaneous reaction of a dislocation with a local stress field and then an interaction with another dislocation

which may be modeled by $C_1 = C_S A$ where A is the in-plane cross sectional area and C_S is the concentration of defects which cause local stress fields at a given thickness.

$$D = \frac{1}{(1/D_0 + C_2/C_1)\exp(C_1 h) - C_2/C_1}, \tag{5.150}$$

where D_0 is a constant. Tachikawa and Yamaguchi used this model to fit their experimental results for GaAs on Si, with $D_0 = 10^{12}\ \text{cm}^{-2}$, $C_1 = 200\ \text{cm}^{-1}$, and $C_2 = 1.8\times10^{-5}$ cm. However, they noted that the dislocation density was inversely proportional to the thickness for all uncracked samples. The cracked samples (thicker than about 10 μm) appeared to exhibit a different thickness dependence, which could be fit by Equation 5.149 but not by the inverse law. However, this behavior has not been observed in other heteroepitaxial material systems for uncracked samples. It is therefore possible that the departure from the inverse law in those samples was related to cracking.

Romanov et al.[165] extended the annihilation and coalescence model of Tachikawa and Yamaguchi to selective area growth and provided a physical analysis of the constants. Here, the starting equation was the same as that given by Tachikawa and Yamaguchi:

$$\frac{dD}{dh} = -C_1 D - C_2 D^2. \tag{5.151}$$

However, it was assumed that the first-order reaction was due to the loss of threading dislocations to sidewalls in the case of selective area epitaxy. The first-order constant was calculated from $C_1 = G/\lambda$, where λ is a length characterizing the travel necessary to reach a mesa sidewall, and G is a geometric factor associated with the inclination of threading dislocations and $G \approx 1$. The second-order constant was calculated from $C_2 = 2G_{r1}$, where r_1 is a characteristic length for the second-order reaction. Romanov et al. wrote the solution in the form

$$D = \frac{D_0}{(1 + C_2 D_0/C_1)\exp\left[C_1(h - h_0)\right] - D_0 C_2/C_1}, \tag{5.152}$$

and discussed two limiting cases. First, for planar (unpatterned) layers, the glide of dislocations to sidewalls will be negligible, so $C_1(h - h_0) \ll 1$ and

$$D = \frac{D_0}{1 + D_0 C_2 (h - h_0)}. \tag{5.153}$$

For large-area growth, the dislocation density will exhibit an (approximately) inverse relationship with the thickness. In the second limiting case, of selective epitaxy with small mesas, the first-order reaction dominates, leading to

$$D = D_0 \exp\left[-C_1(h - h_0)\right]. \tag{5.154}$$

This might explain the different thickness dependence found by Tachikawa and Yamaguchi for thick, cracked films of GaAs/Si(001), which would have contained many sidewalls, but this has not been clarified in a quantitative way.

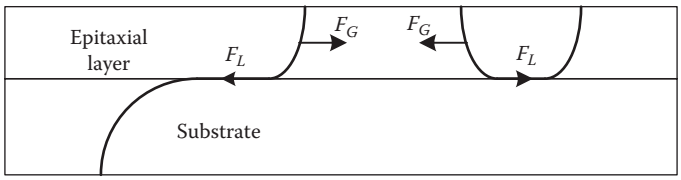


FIGURE 5.67
Forces acting on dislocations in a heteroepitaxial structure. (Adapted from J. E. Ayers, *J. Appl. Phys.*, 78, 3725 [1995]. With permission. Copyright 1995, American Institute of Physics.)

A limitation of the annihilation/coalescence models is that they do not explicitly address the influence of the lattice mismatch.* However, it is known that the dislocation densities in relaxed heteroepitaxial layers exhibit a weak dependence on the lattice mismatch. The "glide model"[166] was developed in an attempt to include the lattice mismatch dependence. Here, it was assumed that reaction (annihilation or coalescence) between two dislocations is limited by their ability to overcome the line tensions of their misfit segments so they can glide toward one another. Figure 5.67 shows the forces acting on dislocations in a relaxed heteroepitaxial layer.† Here, two neighboring threading dislocations have opposite Burgers vectors, resulting in an attractive glide force F_G that acts on each. At the same time, each dislocation experiences a line tension F_L associated with its misfit dislocation segment. If the sample is held at elevated temperature for a sufficiently long time, either during growth or other thermal processing, the dislocations will come to a stable configuration in which the glide and line tension forces balance. Therefore, the dislocation density can be estimated by considering the balance of these forces.

The attractive glide force between dislocations with a separation r is given by

$$F_G = \frac{Gb^2 h}{\pi r \cos\phi}, \tag{5.155}$$

where:
- G is the shear modulus
- b is the length of the Burgers vector
- ϕ is the angle between the threading segments and the interface

The line tension of the misfit segment is given by

$$F_L = \frac{Gb^2}{4\pi}\left[\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right]\ln\left(\frac{R}{4b}\right), \tag{5.156}$$

where:
- α is the angle between the Burgers vector and line vector
- ν is the Poisson ratio
- R is the average spacing between dislocations (perpendicular to the intersection of the glide plane and the heterointerface) or the film thickness, whichever is smaller

* The mismatch dependence is embodied in the parameter D_0.

† Here, the interaction shown in the figure involves one bent-over substrate dislocation and one half loop. Any combination of these types of dislocations can interact, but the stability condition will be the same.

The average spacing for misfit segments is

$$R_{\text{ave}} = \frac{b \sin\alpha \sin\phi}{|f|}, \tag{5.157}$$

Thus,

$$F_L = \frac{Gb^2}{4\pi}\left[\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right]\ln\left(\frac{\sin\alpha\sin\phi}{4|f|}\right). \tag{5.158}$$

We can find the minimum stable separation of threading dislocations by equating the glide and line tension forces, yielding

$$\frac{1}{r_{\min}} = \frac{\cos\phi}{4h}\left[\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right]\ln\left(\frac{\sin\alpha\sin\phi}{4|f|}\right). \tag{5.159}$$

In the development of the glide model, the average value of r was considered to be twice the minimum. (001) heteroepitaxy of a zinc blende semiconductor was assumed, with threading dislocations separated by R_{ave} between glide planes and r_{ave} within glide planes. Using these assumptions, the threading dislocation density was calculated to be

$$D = \frac{2}{R_{\text{ave}} r_{\text{ave}}} = \left[\frac{|f|}{4bh\sin\alpha\tan\phi}\right]\left[\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right]\ln\left(\frac{\sin\alpha\sin\phi}{4|f|}\right). \tag{5.160}$$

Therefore, this model predicts the $1/h$ dependence, and also that the threading dislocation density will increase in sublinear fashion with the lattice mismatch. This model correctly predicts the threading dislocation densities in a number of heteroepitaxial systems, with factor of two accuracy, and without adjustable parameters. However, more experimental results are needed to test this model adequately with respect to the lattice mismatch dependence.

5.8.1 Thermal Strain

Often, heteroepitaxial layers have thermal coefficients of expansion very different from those of their substrates. This causes the introduction of a thermal strain during any temperature ramping process, and in particular during the cooldown to room temperature. This is a severe problem in cases for which the epitaxial layer has a larger thermal coefficient of expansion, because this leads to tensile strain and the possibility of cracking or gross failure of the heteroepitaxial layer. Whereas lattice mismatch strain is often nearly relaxed at the growth temperature, thermal strain is applied during the cooldown and cannot usually be relaxed by dislocation motion. The reason is that dislocation glide velocities are thermally activated and may reduce by a decade for every 25°C reduction in temperature.

Consider a heteroepitaxial layer grown at a temperature T_g and cooled down to room temperature T_r. If the linear coefficients of thermal expansion are α_e and α_s for the epitaxial

layer and substrate, respectively, and if no lattice relaxation occurs during cooldown, then the thermal strain will be

$$\varepsilon_{Th} = \int_{T_g}^{T_r} \left[\alpha_s(T) - \alpha_e(T)\right] dT \,. \tag{5.161}$$

In general, the linear coefficient of thermal expansion is a function of temperature, and this temperature dependence may be fit by an empirical polynomial as described in Chapter 2:

$$\alpha_s = B_s + 2C_s T + 3D_s T^2 \tag{5.162}$$

and

$$\alpha_e = B_e + 2C_e T + 3D_e T^2 \,. \tag{5.163}$$

Therefore, in terms of the thermal expansion polynomials, the thermal strain is given by

$$\varepsilon_{Th} = \left(B_s - B_e\right)\left(T_r - T_g\right) + \left(C_s - C_e\right)\left(T_r^2 - T_g^2\right) + \left(D_s - D_e\right)\left(T_r^3 - T_g^3\right) . \tag{5.164}$$

If the coefficients of thermal expansion are considered to be independent of temperature, then the thermal strain is approximately

$$\varepsilon_{Th} \approx \left(\alpha_e - \alpha_s\right)\left(T_g - T_r\right) . \tag{5.165}$$

The thermal strain is therefore tensile (positive) if the epitaxial layer has a larger coefficient of thermal expansion than its substrate. For a heteroepitaxial layer with a growth temperature in-plane strain of $\varepsilon_{\|}(T_g)$, the room temperature strain will be

$$\varepsilon_{\|}\left(T_r\right) \approx \varepsilon_{\|}\left(T_g\right) + \left(\alpha_e - \alpha_s\right)\left(T_g - T_r\right). \tag{5.166}$$

An interesting practical application of the thermal strain arises in some heteroepitaxial material systems, including certain II-VI semiconductors deposited on GaAs(001) substrates. If the lattice mismatch strain is compressive but the thermal strain is tensile, it is possible to have perfect strain compensation at room temperature. In the case of GaN/sapphire (0001) or GaAs on Si,[139] the material tends to relax almost fully at the growth temperature so that the room temperature strain is always tensile. The same is true for thick layers of ZnSe/GaAs(001).[167] This tensile stress causes warping of large wafers, which increases linearly with the thickness of the epitaxial layer,[168] and could be problematic in device and circuit fabrication. It has been shown, however, that this thermal strain may be eliminated through the application of a strain to the substrate during growth, by mechanical bending.[169]

It has been proposed by Yamaguchi et al.[170] that the high threading dislocations observed in GaAs/Si(001) are associated with the thermal strain. These authors showed that the etch pit density in this material could be diminished considerably by patterning and annealing,

with lateral dimensions on the order of 10 μm. The thermal strain was also reduced in the patterned material, lending support to their model, although another explanation might be glide of dislocations to the sidewalls.

5.8.2 Cracking in Thick Films

Heteroepitaxial films under tensile thermal stress, such as GaN/sapphire (0001), GaN/Si(111), GaAs/Si(001), or InP/Si(001), often develop macroscopic cracks if grown too thick. Figure 5.68 shows an example of cracks in heteroepitaxial AlGaN/sapphire (0001). The simple model of Griffith[171–173] predicts the thickness at which cracks may propagate, for a given tensile strain. This calculation is based on a balance between the strain energy relieved by the crack and the surface energy of the crack walls; a crack may propagate if the film thickness h is equal to or greater than the thickness at which these two energy contributions balance. Films under compressive stress will not crack because cracking would increase their strain energy. They may separate from the substrate, but this is expected to occur at greater thicknesses than cracking, because of the expansive creation of surface area involved in this process.

The Griffith criterion for crack propagation in heteroepitaxial films may be derived as follows.[168–170] Consider a semiconductor layer of thickness h under tensile strain, with a crack passing through the entire thickness and with a length L. The change in surface energy associated with the two walls of the crack will be $2\gamma hL$, where γ is the surface energy per unit area for the semiconductor. At the same time, the crack will relieve strain energy equal to $\pi Lh^2\sigma_{\|}^{\,2} / Y$, where $\sigma_{\|}$ is the biaxial stress and Y is the biaxial modulus. The total change in free energy associated with the crack is

$$\Delta W = 2\gamma hL - \frac{\pi Lh^2\sigma_{\|}^{\,2}}{Y}. \tag{5.167}$$

The condition for propagation of the crack is

$$0 = \frac{\partial(\Delta W)}{\partial L} = 2\gamma h - \frac{\pi h^2\sigma_{\|}^{\,2}}{Y} \tag{5.168}$$

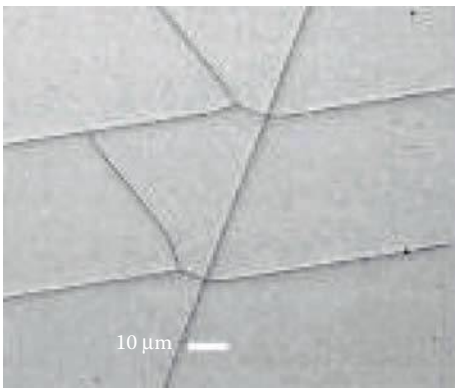


FIGURE 5.68
SEM image of cracks in heteroepitaxial $Al_{0.2}Ga_{0.8}N$ grown on sapphire (0001) by MOVPE, grown with an AlN buffer layer. (Reprinted from J. P. Zhang et al., *Appl. Phys. Lett.*, 80, 3542 [2002]. With permission. Copyright 2002, American Institute of Physics.)

or

$$h = \frac{2Y\gamma}{\pi\sigma_{\|}^{2}}. \tag{5.169}$$

In terms of the strain, the Griffith criterion for crack propagation in a layer with a tensile strain is

$$h \geq \frac{2\gamma}{\pi Y\varepsilon_{\|}^{2}} \qquad \text{(crack propagation)}. \tag{5.170}$$

For (001) heteroepitaxy of a zinc blende semiconductor, the surface energy for the (110) cleavage planes may be approximated by

$$\gamma(110) = \frac{Y(1-\nu)a}{2\sqrt{2}\pi^{2}}, \tag{5.171}$$

where:

- a is the lattice constant
- ν is the Poisson ratio

so that the Griffith criterion becomes

$$h \geq \frac{a(1-\nu)^{2}}{\sqrt{2}\pi^{3}\varepsilon_{\|}^{2}} \quad \text{(crack propagation, zinc blende (001))}. \tag{5.172}$$

Figure 5.69 shows the Griffith thickness as a function of the in-plane tensile strain, for (001) heteroepitaxy of a zinc blende or diamond semiconductor. These calculations show that cracking will be unimportant if the tensile strain is less than about 10^{-4}. On the other hand, an in-plane tensile strain of greater than 1% will result in a Griffith crack thickness less than 0.2 μm and severely limit device design.

The Griffith criterion describes the condition under which existing cracks may propagate, but it is not in itself a sufficient condition for cracking. However, defects or regions of concentrated stress near the edges of a sample may act as nucleation sites for cracks, and we can expect cracks to appear at a thickness close to the Griffith thickness h_G.

The Griffith model as developed above cannot account for cracks with different orientations or irregular geometries. Nonetheless, it could be expected to predict the thickness for the onset of cracking within about a factor of two or three. For example, in the GaAs/Si(001) system, cracking is observed at thicknesses greater than about 4 μm, whereas the Griffith thickness is about 1.5 μm for typical values of the thermal tensile strain.

The thermal strain increases linearly with the growth temperature. Therefore, due to the $1/\varepsilon_{\|}^{2}$ dependence, the Griffith thickness decreases strongly with increasing growth temperature. In GaN/Si(111), for example, GaN can be grown crack-free on Si substrates by MBE at 800°C up to a thickness of 3 μm.[174] But for MOVPE growth, typically carried out at 1090°C, the maximum thickness for crack-free growth is 1.4 μm.[175]

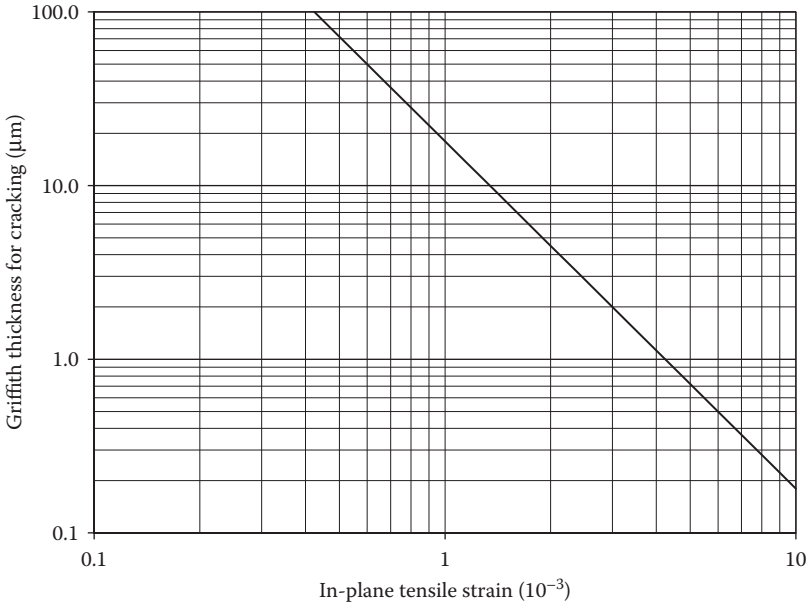


FIGURE 5.69
Griffith thickness for cracking in a heteroepitaxial layer as a function of in-plane tensile strain. (001) heteroepitaxy of a zinc blende or diamond semiconductor was assumed, with $a = 0.565$ nm and $\nu = 1/3$.

For nitride semiconductors grown on Si(111) or sapphire (0001), it has been found that cracking can be suppressed by the insertion of a strained layer superlattice (SLS). For example, Feltin et al.[176] found that they could grow crack-free GaN on Si(111) up to a thickness of 2.5 μm by MOVPE when an AlN/GaN superlattice was inserted (without such a superlattice, the maximum crack-free thickness is about 1.4 μm). The benefit of the SLS was found to come from the introduction of a compressive component of strain in the overlying GaN. For a 0.9 μm thick layer of GaN on Si(111), the room temperature strain without an SLS buffer is about 0.15%, but it can be made negative (compressive) by the insertion of four SLSs. Unfortunately, this effect is diminished with the thickness of the GaN overlayer, and this limits the maximum thickness for crack-free growth.

A similar enhancement in the Griffith thickness was found for MOVPE $Al_{0.2}Ga_{0.8}N$ grown on sapphire (0001) by Zhang et al.,[177] with the insertion of $AlN/Al_{0.2}Ga_{0.8}N$ strained layer superlattices. For $Al_{0.2}Ga_{0.8}N$ top layers grown without the SLS, the maximum crack-free thickness was 1.2 μm, but with an SLS, layers could be grown up to 3.0 μm thick without cracks. As in the previously mentioned study, Zhang et al. found that the insertion of the SLS could compensate the tensile thermal strain and even result in a net compressive strain at room temperature.

PROBLEMS

1. Calculate the critical layer thickness for $In_{0.15}Ga_{0.85}As/GaAs(001)$ assuming (a) 60° misfit dislocations and (b) pure edge misfit dislocations.
2. For $In_xGa_{1-x}As$ grown on InP (001) with $f = 0.4\%$, (a) determine the composition x, (b) calculate the Matthews and Blakeslee critical layer thickness, (c) estimate the

critical layer thickness assuming the surface energy is 2000 erg/cm^2, (d) repeat (a) for $f = -0.4\%$, and (e) repeat (c) for $f = -0.4\%$.

3. For a zinc blende epitaxial layer with (001) orientation, the biaxial modulus is $Y = C_{11} + C_{12} - 2C_{12}^2 / C_{11}$. Show that in the general case, the biaxial modulus for a cubic crystal is

$$Y = \frac{(C_{11} + 2C_{12})}{2}\left[3 - \frac{C_{11} + 2C_{12}}{C_{11} + 2(2C_{44} - C_{11} + C_{12})(l^2m^2 + m^2n^2 + n^2l^2)}\right],$$

where l, m, and n are the direction cosines that relate the unit normal to the cube axes.

4. Suppose that mismatched heteroepitaxy is to be used to bend over all of the dislocations in a substrate having a threading dislocation density of 10^4 cm^{-2}. Find the necessary lattice mismatch. What is the requirement on the thicknss of the heteroepitaxial layer?
5. Suppose a pseudomorphic layer of AlAs is grown on a vicinal (001) GaAs substrate for which the surface normal is inclined by 2° toward the [110]. Find the magnitude and direction for the expected tilt between the heteroepitaxial layer and its substrate.
6. GaN is grown on a vicinal sapphire (0001) substrate that is inclined by 4° toward the $[1\bar{1}00]$. Find the magnitude and direction of the tilt between the GaN and the sapphire, for the case of substrate surface steps with a height of (a) two bilayers, (b) three bilayers, (c) four bilayers, and (d) five bilayers.
7. Find the critical layer thickness for a linearly graded $Si_{1-x}Ge_x$/Si(001) layer if the germanium composition x is graded by 4%/μm. Compare this with the critical layer thickness for a uniform alloy layer having the same germanium surface concentration, if both layers are 0.5 μm thick.
8. (a) Estimate the room temperature thermal strain in InP on Si(001) grown at 650°C. (b) Estimate the thickness beyond which the thermal strain will cause cracking.

References

1. D. K. Ach, J. L. Schmit, R. N. Horning, and J.-L. Staudenmann, Evidence for GaAs substrate strain caused by a CdTe epitaxial layer, *J. Cryst. Growth*, 71, 149 (1985).
2. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers. I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).
3. W. A Jesser and J. W. Matthews, Evidence for pseudomorphic growth of iron on copper, *Phil. Mag.*, 15, 1097 (1967).
4. J. W. Matthews, *Epitaxial Growth, Part B* (New York: Academic Press, 1975).
5. A. Braun, K. M. Briggs, and P. Böni, Analytical solution to Matthews' and Blakeslee's critical dislocation formation thickness of epitaxially grown thin films, *J. Cryst. Growth*, 241, 231 (2002).
6. C. H. Belgacem and M. Fnaiech, Solution for the critical thickness models of dislocation generation in epitaxial thin films using the Lambert W function, *J. Mater. Sci.*, 46, 1913 (2011).
7. J. H. van der Merwe, Crystal interfaces. Part II. Finite overgrowths, *J. Appl. Phys.*, 34, 123 (1962).

8. R. People and J. C. Bean, Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures, *Appl. Phys. Lett.*, 47, 322 (1985); *Appl. Phys. Lett.*, 49, 229 (1986).
9. J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy, *J. Vac. Sci. Technol. A*, 2, 436 (1984).
10. J. Bevk, J. P. Mannaerts, B. A. Davidson, L. C. Feldman, and A. Qurmazd, *Bull. Am. Phys. Soc.*, 31, 533 (1986).
11. D. C. Houghton, C. J. Gibbings, C. G. Tuppen, M. H. Lyons, and M. A. G. Halliwell, Equilibrium critical thickness for $Si_{1-x}Ge_x$ strained layers on (100) Si, *Appl. Phys. Lett.*, 56, 460 (1990).
12. B. Elman, E. S. Koteles, P. Malman, C. Jagannath, J. Lee, and D. Dugger, In situ measurements of critical layer thickness and optical studies of InGaAs quantum wells grown on GaAs substrates, *Appl. Phys. Lett.*, 55, 1659 (1989).
13. M. Gendry and G. Hollinger, Differences in relaxation mechanisms of compressive and tensile strained InGaAs layers grown on InP, 572 (1995).
14. J. Petruzzello and M. R. Leys, Effect of the sign of misfit strain on the dislocation structure at interfaces of heteroepitaxial $GaAs_xP_{1-x}$ films, *Appl. Phys. Lett.*, 53, 2414 (1988).
15. R. C. Cammarata and K. Sieradzki, Surface stress effects on the critical film thickness for epitaxy, *Appl. Phys. Lett.*, 55, 1197 (1989).
16. R. C. Cammarata, K. Sieradzki, and F. Spaepen, Simple model for interface stresses with application to misfit dislocation generation in epitaxial thin films, *J. Appl. Phys.*, 87, 1227 (2000).
17. S. Luryi and E. Suhir, New approach to the high quality epitaxial growth of lattice-mismatched materials, *Appl. Phys. Lett.*, 49, 140 (1986).
18. S. Srinivasan, L. Geng, R. Liu, F. A. Ponce, Y. Narukawa, and S. Tanaka, Slip systems and misfit dislocations in InGaN epilayers, *Appl. Phys. Lett.*, 83, 5187 (2003).
19. D. Holec, P. M. F. J. Costa, M. J. Kappers, and C. J. Humphreys, Critical thickness calcualtions for InGaN/GaN, *J. Cryst. Growth*, 303, 314 (2007).
20. D. Holec, Y. Zhang, D. V. S. Rao, M. J. Kappers, and C. McAleese, Equilibrium critical thickness for misfit dislocations in III-nitrides, *J. Appl. Phys.*, 104, 123514 (2008).
21. H. Ye, P. Lu, Z. Yu, Z. Chen, B. Jia, H. Feng, and Y. Liu, Equilibrium critical thickness for a wurtzite InGaN/GaN heterostructure, *Superlat. Microstruct.*, 48, 58 (2010).
22. I. Belabbas, A. Bere, J. Chen, S. Petit, M. A. Belkhir, P. Ruterana, and G. Nouet, Atomistic modeling of the (a+c)-mixed dislocation core in wurtzite GaN, *Phys. Rev. B*, 75, 115201 (2007).
23. L. Lymperakis, Ab-initio based multiscale calculations of extended defects in and on group III-nitrides, PhD thesis, Department Physik der Fakultat fur Naturwissenschaften an der Universitat Paderborn (2005).
24. J. Kioseoglou, Ph. Komninou, and Th. Karakostas, Core models of a-edge threading dislocations in wurtzite III(Al,Ga,In)-nitrides, *Phys. Status Solidi A*, 206, 1931 (2009).
25. J. A. Floro, D. M. Follstaedt, P. Provencio, S. J. Hearne, and S. R. Lee, Misfit dislocation formation in the AlGaN/GaN heterointerface, *J. Appl. Phys.*, 96, 7087 (2004).
26. S. R. Lee, D. D. Koleske, K. C. Cross, J. A. Floro, K. E. Waldrip, A. T. Wise, and S. Mahajan, In situ measurements of the critical layer thickness for strain relaxation in AlGaN/GaN heterostructures, *Appl. Phys. Lett.*, 85, 6164 (2004).
27. P. Vennegues, Z. Bougrioua, J. M. Bethoux, M. Azize, and O. Tottereau, Relaxation mechanisms in metal-organic vapor phase epitaxy grown Al-rich (Al,Ga)N/GaN heterostructures, *J. Appl. Phys.*, 97, 024912 (2005).
28. J.-M. Bethoux and P. Vennegues, Ductile relaxation in cracked metal-organic chemical-vapor-deposition-grown AlGaN films on GaN, *J. Appl. Phys.*, 97, 123504 (2005).
29. M. Gherasimova, G. Cui, Z. Ren, J. Su, X.-L. Wang, J. Han, K. Higashimine, and N. Otsuka, Heteroepitaxial evolution of AlN on GaN grown by metal-organic chemical vapor deposition, *J. Appl. Phys.*, 95, 2921 (2004).
30. R. Liu, J. Mei, S. Srinivasan, F. A. Ponce, H. Omiya, Y. Narukawa, and T. Mukai, Generation of misfit dislocations by basal-plane slip in InGaN/GaN heterostructures, *Appl. Phys. Lett.*, 89, 201911 (2006).

31. W. Lu, D. B. Li, C. R. Li, and Z. Zhang, Generation and behavior of pure-edge threading dislocations in $In_xGa_{1-x}N$/GaN multiple quantum wells, *J. Appl. Phys.*, 5267 (2004).
32. R. Liu, J. Mei, S. Srinivasan, H. Omiya, F. A. Ponce, D. Cherns, Y. Narukawa, and T. Mukai, Misfit dislocation generation in InGaN epilayers on free-standing GaN, *Jpn. J. Appl. Phys.*, 45, L549 (2006).
33. H. K. Cho and G. M. Yang, Generation of misfit dislocations in high indium content InGaN layer grown on GaN, *J. Cryst. Growth*, 243, 124 (2002).
34. B. Jahnen, M. Albrecht, W. Dorsch, S. Christiansen, H. P. Strunk, D. Hanser, and R. F. Davis, Pinholes, dislocations, and strain relaxation in InGaN, *MRS Internet J. Nitride Semicond.*, 3, 39 (1998).
35. P. M. F. J. Costa, R. Datta, M. J. Kappers, M. E. Vickers, C. J. Humphreys, D. M. Graham, P. Dawson, M. J. Godfrey, E. J. Thrush, and J. T. Mullins, Misfit dislocations in In-rich InGaN/GaN quantum well structures, *Phys. Status Solidi A*, 203, 1729 (2006).
36. C. A. Parker, J. C. Roberts, S. M. Bedair, M. J. Reed, S. X. Liu, and N. A. El-Masry, Determination of the critical layer thickness in the InGaN/GaN heterostructures, *Appl. Phys. Lett.*, 75, 2776 (1999).
37. M. J. Reed, N. A. El-Masry, C. A. Parker, J. C. Roberts, and S. M. Bedair, Critical layer thickness determination of GaN/InGaN/GaN double heterostructures, *Appl. Phys. Lett.*, 77, 4121 (2000).
38. J. Kioseoglou, E. Kalesaki, L. Lymperakis, G. P. Dimititrakopoulos, Ph. Komninou, and Th. Karakostas, Polar AlN/GaN interface: Structures and energetics, *Phys. Status Solidi A*, 206, 1892 (2009).
39. J. W. Matthews, Defects associated with the accommodation of misfit between crystals, *J. Vac. Sci. Technol.*, 12, 126 (1975).
40 J. W. Matthews, A. E. Blakeslee, and S. Mader, Use of misfit strain to remove dislocations from epitaxial thin films, *Thin Solid Films*, 33, 253 (1976).
41. B. W. Dodson, Nature of misfit dislocation sources in strained-layer semiconductor structures, *Appl. Phys. Lett.*, 53, 394 (1988).
42. H. Gao, C. S. Ozkan, W. D. Nix, J. A. Zimmerman, and L. B. Freund, Atomistic models of dislocation formation at crystal surface ledges in $Si_{1-x}Ge_x$/Si(100) heteroepitaxial thin films, *Phil. Mag. A*, 79, 349 (1999).
43. J. Zou, X. Z. Liao, D. J. H. Cockayne, and Z. M. Jiang, Alternative mechanism for misfit dislocation generation during high-temperature Ge(Si)/Si(001) island growth, *Appl. Phys. Lett.*, 81, 1996 (2002).
44. R. Beanland, Multiplication of misfit dislocations in epitaxial layers, *J. Appl. Phys.*, 72, 4031 (1992).
45. J. W. Matthews, S. Mader, and T. B. Light, Accommodation of misfit across the interface between crystals of semiconducting elements or compounds, *J. Appl. Phys.*, 41, 3800 (1970).
46. F. C. Frank and W. T. Read, Multiplication processes for slow moving dislocations, *Phys. Rev.*, 79, 722 (1950).
47. W. C. Dash, Copper precipitation on dislocations in silicon, *J. Appl. Phys.*, 27, 1193 (1956).
48. A. T. Lefevbre, C. Herbeaux, C. Bouillet, and J. Di Persio, A new type of misfit dislocation multiplication process in $In_xGa_{1-x}As$/GaAs strained-layer superlattices, *Phil. Mag. Lett.*, 63, 23 (1991).
49. F. K. LeGoues, B. S. Meyerson, and J. M. Morar, Anomalous strain relaxation in SiGe thin films and superlattices, *Phys. Rev. Lett.*, 66, 2903 (1991).
50. M. A. Capano, L. Hart, D. K. Bowen, D. Gordon-Smith, C. R. Thomas, C. J. Gibbings, M. A. G. Halliwell, and L. W. Hobbs, Strain relaxation in $Si_{1-x}Ge_x$ layers on Si(001), *J. Cryst. Growth*, 116, 260 (1992).
51. W. C. Dash, *Dislocations and Mechanical Properties of Crystals*, ed. J. Fisher (New York: John Wiley, 1957), p. 57.
52. A. Authier and A. R. Lang, Three-dimensional x-ray topographic studies of internal dislocation sources in silicon, *J. Appl. Phys.*, 35, 1956 (1964).
53. A. T. Mader and A. E. Blakeslee, On dislocations in $GaAs_{1-x}P_x$, *IBM J. Res. Dev.*, 19, 151 (1985).

54. A. T. Washburn and E. P. Kvam, Possible dislocation multiplication source in (001) semiconductor epitaxy, *Appl. Phys. Lett.*, 57, 1637 (1990).
55. W. Hagen and H. Strunk, A new type of source generating misfit dislocations, *Appl. Phys.*, 17, 85 (1978).
56. H. Strunk, W. Hagen, and E. Bauser, Low-density dislocation arrays at heteroepitaxial Ge/GaAs interfaces investigated by high voltage electron microscopy, *Appl. Phys. A*, 18, 67 (1979).
57. R. H. Dixon and P. J. Goodhew, On the origin of misfit dislocations in InGaAs/GaAs strained layers, *J. Appl. Phys.*, 68, 3163 (1990).
58. Y. Obayashi and K. Shintani, Is the Hagen-Strunk multiplication mechanism of misfit dislocations in heteroepitaxial layers probable? *Phil. Mag. Lett.*, 76, 1 (1997).
59. M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, Dislocation morphology in graded heterojunctions: $GaAs_{1-x}P_x$, *J. Mater. Sci.*, 4, 223 (1969).
60. G. R. Booker, J. M. Titchmarsh, J. Fletcher, D. B. Darby, M. Hockly, and M. Al-Jassim, Nature, origin and effect of dislocations in epitaxial semiconductor layers, *J. Cryst. Growth*, 45, 407 (1978).
61. V. I. Vdovin, L. A. Matveeva, G. N. Semenova, M. Ya. Skorohod, Yu. A. Tkhorik, and L. S. Khazan, Misfit dislocations in epitaxial heterostructures: Mechanisms of generation and multiplication, *Phys. Status Solidi A*, 92, 379 (1985).
62. F. R. Chien, X. J. Ning, S. Stemmer, P. Pirouz, M. D. Bremser, and R. F. Davis, Growth defects in GaN films on 6H-SiC substrates, *Appl. Phys. Lett.*, 68, 2678 (1996).
63. W. Qian, M. Skowronski, and R. Kaspi, Dislocation density reduction in GaSb films grown on GaAs substrates by molecular beam epitaxy, *J. Electrochem. Soc.*, 144, 1430 (1997).
64. W. Qian, M. Skowronski, R. Kaspi, and M. De Graef, Nucleation of misfit and threading dislocations during epitaxial growth, *J. Appl. Phys.*, 81, 7268 (1997).
65. P. M. J. Marée, R. I. J. Olthof, J. W. M. Frenken, J. F. van der Veen, C. W. T. Bulle-Lieuwma, M. P. A. Viegers, and P. C. Zalm, Silicon strained layers grown on GaP(001) by molecular beam epitaxy, *J. Appl. Phys.*, 58, 3097 (1985).
66. K. Ito, K. Hiramatsu, H. Amano, and I. Akasaki, Preparation of $Al_xGa_{1-x}N$/GaN heterostructure by MOVPE, *J. Cryst. Growth*, 104, 533 (1990).
67. R. E. Mallard, P. R. Wilshaw, N. J. Mason, P. J. Walker, and G. R. Booker, Lattice-relaxation of strained GaSb GaAs epitaxial layers grown by MOCVD, *Inst. Phys. Conf. Ser.*, 100, 331 (1989).
68. G. R. Johnson, B. C. Cavenett, T. M. Kerr, P. B. Kirby, and C. E. C. Wood, Optical, Hall, and cyclotron resonance measurements of GaSb grown by molecular beam epitaxy, *Semicond. Sci. Technol.*, 3, 1157 (1988).
69. S. V. Ivanov, P. D. Altukhov, T. S. Argunova, A. A. Bakun, A. A. Budza, V. V. Chaldyshev, Yu. A. Kovalenko, et al., Molecular beam epitaxy growth and characterization of thin (<2 μm) GaSb layers on GaAs(100) substrates, *Semicond. Sci. Technol.*, 8, 347 (1993).
70. R. N. Kyutt, R. Scholz, S. S. Ruvimov, T. S. Argunova, A. A. Budza, S. V. Ivanov, P. S. Kop'ev, L. M. Sorokin, and M. P. Sheheglov, Dislocation structure of epitaxial GaSb films grown on (001) GaAs substrates by molecular beam epitaxy, *Phys. Solid State*, 35, 372 (1993).
71. J. M. Kang, M. Nouaoura, L. Lassabatere, and A. Rocher, Accommodation of lattice mismatch and threading dislocations in GaSb films grown at different temperatures on GaAs(001), *J. Cryst. Growth*, 143, 115 (1994).
72. J. M. Kang, S.-K. Min, and A. Rocher, Asymmetric tilt interface induced by 60° misfit dislocation arrays in GaSb/GaAs(001), *Appl. Phys. Lett.*, 65, 2954 (1994).
73. P. M. Thibado, B. R. Bennett, M. E. Twigg, B. V. Shanabrook, and L. J. Whitman, Evolution of GaSb epitaxy on GaAs(001)-c(4 × 4), *J. Vac. Sci. Technol. A*, 14, 885 (1996).
74. S. H. Huang, G. Balakrishnan, A. Khoshaklagh, A. Jallipalli, L. R. Dawson, and D. L. Huffaker, Strain relief by periodic misfit arrays for low defect density GaSb on GaAs, *Appl. Phys. Lett.*, 88, 131911 (2006).
75. W. Zhou, X. Li, S. Xia, J. Yang, W. Tang, and K. M. Lau, High hole mobility of GaSb relaxed epilayer grown on GaAs substrate by MOCVD through interfacial misfit dislocations array, *J. Mater. Sci. Technol.*, 28, 132 (2012).

76. J. Tetebayashi, A. Khoshakhlagh, S. H. Huang, L. R. Dawson, G. Balakrishnan, and D. L. Huffaker, Formation and optical characteristics of strain-relieved and densely stacked GaSb/GaAs quantum dots, *Appl. Phys. Lett.*, 89, 203116 (2006).
77. S. H. Huang, G. Balakrishnan, M. Mehta, A. Khoshaklagh, L. R. Dawson, and D. L. Huffaker, Epitaxial growth and formation of interfacial misfit array for tensile GaAs on GaSb, *Appl. Phys. Lett.*, 90, 161902 (2007).
78. S. H. Huang, G. Balakrishnan, A. Khsohakhlagh, L. R. Dawson, and D. L. Huffaker, Simultaneous interfacial misfit array formation and antiphase domain suppression on miscut silicon substrates, *Appl. Phys. Lett.*, 93, 071102 (2008).
79. M. Mehta, G. Balakrishnan, S. Huang, A. Khoshakhlagh, A. Jallipalli, P. Patel, M. N. Kuty, L. R. Dawson, and D. L. Huffaker, GaSb quantum-well-based 'buffer-free' vertical light emitting diode monolithically embedded within a GaAs cavity incorporating interfacial misfit arrays, *Appl. Phys. Lett.*, 89, 211110 (2006).
80. K. G. Eyink, L. Grazulis, J. J. Pitz, J. Shank, and K. Mahalingam, Strain relaxation in the growth of planar InAs, *J. Vac. Sci. Technol. B*, 24, 1581 (2006).
81. C. J. K. Richardson, L. He, and S. Kanakaraju, Metamorphic growth of III-V semiconductor bicrystals, *J. Vac. Sci. Technol. B*, 29, 03C126 (2011).
82. V. M. Kaganer, R. Kohler, M. Schmidbauer, R. Opitz, and B. Jenichen, X-ray diffraction peaks due to misfit dislocations in heteroepitaxial structures, *Phys. Rev. B*, 55, 1793 (1997).
83. R. Hull and J. C. Bean, Experimental and theoretical analysis of strain relaxation in Ge_xSi_{1-x}/Si(100) heteroepitaxy, *Mater. Res. Soc. Symp. Proc.*, 148, 309 (1989).
84. R. Hull, J. C. Bean, and C. Buescher, A phenomenological description of strain relaxation in Ge_xSi_{1-x}/Si(100) heterostructures, *J. Appl. Phys.*, 66, 5837 (1989).
85. B. W. Dodson and J. Y. Tsao, Stress dependence of dislocation glide activation energy in single-crystal silicon-germanium alloys up to 2.6 GPa, *Phys. Rev. B*, 38, 12383 (1988).
86. S. K. Choi, M. Mihara, and T. Ninomiya, Dislocation velocities in GaAs, *Jpn. J. Appl. Phys.*, 16, 737 (1977).
87. S. A. Erofeeva and Yu. A. Osip'yan, Mobility of dislocations in crystals with the sphalerite lattice, *Sov. Phys. Solid State* (Engl. transl.), 15, 538 (1973).
88. J. J. Gilman, Dislocation mobility in crystals, *J. Appl. Phys.*, 36, 3195 (1965).
89. A. R. Chaudhuri, J. R. Patel, and L. G. Rubin, Velocities and densities of dislocations in germanium and other semiconductor crystals, *J. Appl. Phys.*, 33, 2736 (1962).
90. R. Hull and J. C. Bean, Kinetic barriers to strain relaxation in Ge(x)Si(1–x) epitaxy, *Mater. Res. Symp. Proc.*, 160, 23 (1990).
91. A. George and J. Rabier, Dislocations and plasticity in semiconductors. I. Dislocation structures and dynamics, *Rev. Phys. Appl.*, 22, 941 (1987).
92. J. P. Hirth and J. Lothe, *Theory of Dislocations* (New York: McGraw-Hill, 1968).
93. A. Seeger and P. Schiller, Bildung und Diffusion von Kinken als Grundprozess der Versetzungsbewegung bei der Messung der inneren Reibung, ***Acta Metall.***, 10, 348 (1962).
94. R. Hull, J. C. Bean, D. Bahnck, L. J. Peticolas Jr., K. T. Short, and F. C. Unterwald, Interpretation of dislocation propagation velocities in strained Ge_xSi_{1-x}/Si(100) heterostructures by the diffusive kink pair model, *J. Appl. Phys.*, 70, 2052 (1991).
95. L. B. Freund, A criterion for arrest of a threading dislocation in a strained epitaxial layer due to an interface misfit dislocation in its path, *J. Appl. Phys.*, 68, 2073 (1990).
96. D. C. Paine, D. J. Howard, D. Luo, R. N. Sacks, and T. C. Eschrich, The relaxation of $In_xGa_{1-x}As$/GaAs strained multilayers, *Mater. Res. Soc. Symp. Proc.*, 160, 123 (1990).
97. R. Hull and J. C. Bean, Variation in misfit behavior as a function of strain in the GeSi/Si system, *Appl. Phys. Lett.*, 54, 925 (1989).
98. R. Hull, H. Parvaneh, D. Andersen, and J. C. Bean, Materials genomics of thin film strain relaxation by misfit dislocations, *J. Appl. Phys.*, 118, 225306 (2015).
99. V. T. Gillard, W. D. Nix, and L. B. Freund, Role of dislocation blocking in limiting strain relaxation in heteroepitaxial films, *J. Appl. Phys.*, 76, 7280 (1994).

100. C. A. Volkert, E. A. Fitzgerald, R. Hull, Y.-H. Xie, and Y.-J. Mii, Strain relaxation in $Ge_{0.09}Si_{0.91}$ epitaxial thin films measured by wafer curvature, *J. Electron. Mater.*, 20, 833 (1991).
101. R. J. Hauenstein, B. M. Clemens, R. H. Miles, O. J. Marsh, E. T. Croke, and T. C. McGill, Strain relaxation kinetics in $Si_{1-x}Ge_x$/Si heterostructures, *J. Vac. Sci. Technol. B*, 7, 767 (1989).
102. S. B. Samavedam and E. A. Fitzgerald, Novel dislocation structure and surface morphology effects in relaxed Ge/Si-Ge(graqded)/Si structures, *J. Appl. Phys.*, 81, 3108 (1997).
103. P. Haasen, On the plasticity of germanium and indium antimonide, *Acta. Met.*, 5, 598 (1957).
104. B. W. Dodson and J. Y. Tsao, Relaxation of strained-layer semiconductor structures via plastic flow, *Appl. Phys. Lett.*, 51, 1325 (1987); *Appl. Phys. Lett.*, 52, 852 (1988).
105. B. Yarlagadda, A. Rodriguez, P. Li, B. I. Miller, F. C. Jain, and J. E. Ayers, Elastic strains in heteroepitaxial $ZnSe_{1-x}Te_x$ on InGaAs/InP (001), *J. Electron. Mater.*, 35, 1327 (2006).
106. B. W. Dodson, Letter to the editors: An empirical relation for semiconductor strained-layer relaxation, *J. Cryst. Growth*, 91, 649 (1988).
107. D. E. Grider, S. E. Swirhun, D. H. Narum, A. I. Akinwande, T. E. Nohava, W. R. Stuart, P. Joslyn, and K. C. Hsieh, Metamorphic $In_yGa_{1-y}As/In_zAl_{1-z}As$ heterostructure field effect transistors grown on GaAs(001) substrates using molecular-beam epitaxy, *J. Vac. Sci. Technol. B*, 8, 301 (1990).
108. T. Kujofsa, W. Yu, S. Cheruku, B. Outlaw, S. Xhurxhi, F. Obst, D. Sidoti, et al., Plastic flow and dislocation compensation in ZnS_ySe_{1-y}/GaAs(001) heterostructures, *J. Electron. Mater.*, 41, 2993 (2013).
109. T. Kujofsa, S. Cheruku, W. Yu, B. Outlaw, S. Xhurxhi, F. Obst, D. Sidoti, et al., Relaxation dynamics and threading dislocations in ZnS_ySe_{1-y}/GaAs(001) heterostructures, *J. Electron. Mater.*, 42, 2764 (2013).
110. T. Kujofsa and J. E. Ayers, Evolution of kinetically-limited lattice relaxation and threading dislocations in temperature-graded ZnSe/GaAs(001) metamorphic heterostructures, *J. Electron. Mater.*, 44, 3030 (2015).
111. B. K. Tanner, P. J. Parbrook, B. Lunn, J. H. C. Hogg, A. M. Keir, and A. D. Johnson, Observation of hardening during relaxation of InGaAs on GaAs, *J. Phys. D Appl. Phys.*, 36, A198 (2003).
112. P. J. Orderes and B. F. Usher, Determination of critical layer thickness in $In_xGa_{1-x}A$/GaAs heterostructures by x-ray diffraction, *Appl. Phys. Lett.*, 50, 980 (1987).
113. P. Maigné and J. M. Baribeau, Measurement of residual strain in InGaAs buffer layers, *J. Appl. Phys.*, 76, 1962 (1994).
114. S. K. Mathis, P. Chavarkar, A. M. Andrews, U. K. Mishra, and J. S. Speck, Strain relaxation of InGaAs by lateral oxidation of AlAs, *J. Vac. Sci. Technol. B*, 18, 2066 (2000).
115. H. Suzuki, T. Sasaki, A. Sai, Y. Ohshita, I. Kamiya, M. Yamaguchi, M. Takahasi, and S. Fujikawa, Real-time observation of anisotropic strain relaxation by three-dimensional reciprocal space mapping during InGaAs/GaAs(001) growth, *Appl. Phys. Lett.*, 97, 041906 (2010).
116. G. J. Whaley and P. I. Cohen, Relaxation of strained InGaAs during molecular beam epitaxy, *Appl. Phys. Lett.*, 57, 144 (1990).
117. I. Yonenaga, Dynamic behavior of dislocations in InAs: In comparison with III–V compounds and other semiconductors, *J. Appl. Phys.*, 84, 4209 (1998).
118. I. Yonenaga and K. Sumino, Effects of In impurity on the dynamic behavior of dislocations in GaAs, *J. Appl. Phys.*, 62, 1212 (1987).
119. R. Beresford, J. Yin, K. Tetz, and E. Chason, Real-time measurements of stress relaxation in InGaAs/GaAs, *J. Vac. Sci. Technol. B*, 18, 1431 (2000).
120. T. Sasaki, H. Suzuki, A. Sai, M. Takahasi, S. Fujikawa, I. Kamiya, Y. Ohshita, and M. Yamaguchi, Growth temperature dependence of strain relaxation during InGaAs/GaAs(001) heteroepitaxy, *J. Cryst. Growth*, 323, 13 (2011).
121. D. J. Dunstan, Strain and strain relaxation in semiconductors, *J. Mater. Sci. Mater. Electron.*, 8, 337 (1997).
122. C. Lynch, E. Chason, R. Beresford, E. B. Chen, and D. C. Paine, Dislocation structure and relaxation kinetics in InGaAs/GaAs heteroepitaxy, *J. Vac. Sci. Technol. B*, 20, 1247 (2002).

123. T. Sasaki, K. Shimomura, H. Suziki, M. Takahasi, I. Kamiya, O. Ohshita, and M. Yamaguchi, Observation of in-plane asymmetric strain relaxation during crystal growth and growth interruption in InGaAs/GaAs(001), *Jpn. J. Appl. Phys.*, 51, 02BP01 (2012).
124. D. J. Dunstan, P. Kidd, L. K. Howard, and R. H. Dixon, Plastic relaxation of InGaAs grown on GaAs, *Appl. Phys. Lett.*, 59, 3390 (1991).
125. D. J. Dunstan, P. Kidd, P. F. Fewster, N. L. Andrew, R. Grey, J. P. R. David, L. González, Y. González, A. Sacedón, and F. González-Sanz, Plastic relaxation of metamorphic single layer and multilayer InGaAs/GaAs structures, *Appl. Phys. Lett.*, 65, 839 (1994).
126. V. Bellani, C. Bocchi, T. Ciabattoni, S. Franchi, P. Frigeri, P. Galinetto, M. Geddo, et al., Residual strain measurements in InGaAs metamorphic buffer layers on GaAs, *Eur. Phys. J. B*, 56, 217 (2007).
127. V. Krishnamoorthy, Y. W. Lin, L. Calhoun, H. L. Liu, and R. M. Park, Residual strain analysis of $In_xGa_{1-x}As$/GaAs heteroepitaxial layers, *Appl. Phys. Lett.*, 61, 2680 (1992).
128. T. L. Song, S. J. Chua, E. A. Fitzgerald, P. Chen, and S. Tripathy, Plastic relaxation in single $In_xGa_{1-x}N$/GaN epilayers grown on sapphire, *Adv. Mater. Micro Nano Syst*, 2004.
129. X. R. Huang, J. Bai, M. Dudley, R. D. Dupuis, and U. Chowdhury, Epitaxial tilting of GaN grown on vicinal surfaces of sapphire, *Appl. Phys. Lett.*, 86, 211916 (2005).
130. X. R. Huang, J. Bai, M. Dudle, B. Wagner, R. F. Davis, and Y. Zhu, Step-controlled strain relaxation in the vicinal surface epitaxy of nitrides, *Phys. Rev. Lett.*, 95, 86101 (2005).
131. A. Leiberich and J. Levkoff, A double crystal x-ray diffraction characterization of $Al_xGa_{1-x}As$ grown on an offcut GaAs(100) substrate, *J. Vac. Sci. Technol. B*, 8, 422 (1990).
132. H. Nagai, Structure of vapor-deposited $Ga_xIn_{1-x}As$ crystals, *J. Appl. Phys.*, 45, 3789 (1974).
133. G. H. Olsen and R. T. Smith, *Phys. Status Solidi A*, 31, 739 (1975).
134. A. Ohki, N. Shibata, and S. Zembutsu, Lattice relaxation mechanism of ZnSe layer grown on a (100) GaAs substrate tilted toward <011>, *J. Appl. Phys.*, 64, 694 (1988).
135. J. Kleiman, R. M. Park, and H. A. Mar, On epilayer tilt in ZnSe/Ge heterostructures prepared by molecular-beam epitaxy, *J. Appl. Phys.*, 64, 1201 (1988).
136. E. Yamaguchi, I. Takayasu, T. Minato, and M. Kawashima, Growth of ZnSe on Ge(100) substrates by molecular-beam epitaxy, *J. Appl. Phys.*, 62, 885 (1987).
137. I. B. Bhat, K. Patel, N. R. Taskar, J. E. Ayers, and S. K. Ghandhi, X-ray diffraction studies of CdTe grown on InSb, *J. Cryst. Growth*, 88, 23 (1988).
138. W. L. Ahlgren, S. M. Johnson, E. J. Smith, R. P. Ruth, B. C. Johnston, M. H. Kalisher, C. A. Cokrum, T. W. James, D. L. Arney, C. K. Ziegler, and W. Lick, Metalorganic chemical vapor deposition growth of $Cd_{1-y}Zn_yTe$ epitaxial layers on GaAs and GaAs/Si, *J. Vac. Sci. Technol. A*, 7, 331 (1989).
139. T. J. de Lyon, D. Rajavel, S. M. Johnson, and C. A. Cockrum, Molecular-beam epitaxial growth of CdTe(112) on Si(112) substrates, *Appl. Phys. Lett.*, 66, 2119 (1995).
140. J. W. Lee, J. P. Salerno, R. P. Gale, and J. C. C. Fan, Epitaxy of GaAs on Si: MBE and OMCVD, *Mater. Res. Soc. Symp. Proc.*, 91, 33 (1987).
141. L. J. Schowalter, E. L. Hall, N. Lewis, and S. Hashimoto, Strain relief of large lattice mismatch heteroepitaxial films on silicon by tilting, *Thin Solid Films*, 184, 437 (1990).
142. S. K. Ghandhi and J. E. Ayers, Strain and misorientation in GaAs grown on Si(001) by organometallic epitaxy, *Appl. Phys. Lett.*, 53, 1204 (1988).
143. F. Riesz, J. Varrio, A. Pesek, and K. Lischka, Correlation between initial growth planarity and epilayer tilting in the vicinal GaAs/Si system, *Appl. Surface Sci.*, 75, 248 (1994).
144. M. Calamiotou, N. Chrysanthakopoulos, Ch. Lioutas, K. Tsagaraki, and A. Georgakilas, Microstructural differences of the two possible orientations of GaAs on vicinal (001) Si substrates, *J. Cryst. Growth*, 227–228, 98 (2001).
145. Y.-Z. Yoo, T. Chikyow, M. Kawasaki, T. Onuma, S. Chichibu, and H. Koinuma, Heteroepitaxy of hexagonal ZnS thin films directly on Si(111), *Jpn. J. Appl. Phys.*, Part 1, 42, 7029 (2003)
146. X. Jiang, R. Q. Zhang, G. Yu, and S. T. Lee, Local strain in interface: Origin of grain tilting in diamond (001)/silicon (001) heteroepitaxy, *Phys. Rev. B*, 58, 15351 (1998).

147. X. Jiang, K. Schiffmann, C.-P. Klages, D. Wittorf, C. L. Jia, K. Urban, and W. Jager, Coalescence and overgrowth of diamond grains for improved heteroepitaxy on silicon (001), *J. Appl. Phys.*, 83, 2511 (1998).
148. E. Maillard-Schaller, O. M. Kuttel, P. Groning, O. Groning, R. G. Agostino, P. Aebi, and L. Schlapbach, Local heteroepitaxy of diamond on silicon (100): A study of the interface structure, *Phys. Rev. B*, 55, 15895 (1997).
149. B. W. Dodson, D. R. Meyers, A. K. Datye, V. S. Kaushik, D. L. Kendall, and B. Martinez-Tovar, Asymmetric tilt boundaries and generalized heteroepitaxy, *Phys. Rev. Lett.*, 61, 2681 (1988).
150. P. Auvray, M. Baudet, and A. Regreny, X-ray diffraction effects in Ga and Al arsenide structures MBE-grown on slightly misoriented GaAs(001) substrates, *J. Cryst. Growth*, 95, 288 (1989).
151. J. E. Ayers, S. K. Ghandhi, and L. J. Schowalter, Crystallographic tilting of heteroepitaxial layers, *J. Cryst. Growth*, 113, 430 (1991).
152. J. E. Ayers and L. J. Schowalter, Comment on "Measurement of the activation barrier to nucleation of dislocations in thin films," *Phys. Rev. Lett.*, 72, 4055 (1994).
153. J. Y. Tsao and B. W. Dodson, Excess stress and the stability of strained heterostructures, *Appl. Phys. Lett.*, 53, 848 (1988).
154. F. Riesz, Crystallographic tilting in lattice-mismatched heteroepitaxy: A Dodson-Tsao relaxation approach, *J. Appl. Phys.*, 79, 4111 (1996).
155. T. W. Weeks Jr., M. D. Bremser, K. S. Ailey, E. Carlson, W. G. Perry, and R. F. Davis, GaN thin films deposited via organometallic vapor phase epitaxy on α(6H)-SiC(0001) using high-temperature monocrystalline AlN buffer layers, *Appl. Phys. Lett.*, 67, 401 (1995).
156. M. H. Xie, L. X. Zheng, S. H. Cheung, Y. F. Ng, H. Wu, S. Y. Tong, and N. Ohtani, Reduction of threading defects in GaN grown on vicinal SiC (0001) by molecular-beam epitaxy, *Appl. Phys. Lett.*, 77, 1105 (2000).
157. J.-I. Kato, S. Tanaka, S. Yamada, and I. Suemune, Structural anisotropy in GaN films grown on vicinal 4H-SiC surfaces by metalorganic molecular-beam epitaxy, *Appl. Phys. Lett.*, 83, 1569 (2003).
158. L. Lu, H. Yan, C. L. Yang, M. Xie, Z. Wang, J. Wang, and W. Ge, Study of GaN thin films grown on vicinal SiC(0001) substrates by molecular beam epitaxy, *Semicond. Sci. Technol.*, 17, 957 (2002).
159. X.-Q. Shen, M. Shimizu, and H. Okumura, Impact of vicinal sapphire (0001) substrates on the high-quality AlN films by plasma-assisted molecular beam epitaxy, *Jpn. J. Appl. Phys.*, 42, L1293 (2003).
160. L. P. Van, O. Kurnosikov, and J. Cousty, Evolution of steps on vicinal (0001) surfaces of α-alumina, *Surf. Sci.*, 411, 263 (1998).
161. P. Sheldon, K. M. Jones, M. M. Al-Jassim, and B. G. Yacobi, Dislocation density reduction through annihilation in lattice-mismatched semiconductors grown by molecular beam epitaxy, *J. Appl. Phys.*, 63, 5609 (1988).
162. J. E. Ayers, L. J. Schowalter, and S. K. Ghandhi, Post-growth thermal annealing of GaAs on Si(001) grown by organometallic vapor phase epitaxy, *J. Cryst. Growth*, 125, 329 (1992).
163. S. Akram, H. Ehsani, and I. B. Bhat, The effect of GaAs surface stabilization on the properties of ZnSe grown by organometallic vapor phase epitaxy, *J. Cryst. Growth*, 124, 628 (1992).
164. S. Kalisetty, M. Gokhale, K. Bao, J. E. Ayers, and F. C. Jain, The influence of impurities on the dislocation behavior in heteroepitaxial ZnSe on GaAs, *Appl. Phys. Lett.*, 68, 1693 (1996).
165. A. E. Romanov, W. Pompe, G. E. Beltz, and J. S. Speck, An approach to threading dislocation 'reaction kinetics,' *Appl. Phys. Lett.*, 69, 3342 (1996).
166. J. E. Ayers, New model for the thickness and mismatch dependencies of threading dislocation densities in mismatched heteroepitaxial layers, *J. Appl. Phys.*, 78, 3724 (1995).
167. B. J. Skromme, M. C. Tamargo, F. S. Turco, S. M. Shibli, W. A. Bonner, and R. E. Nahory, Characterization of heteroepitaxial ZnSe grown by MBE on GaAs, AlAs, and InGaAs, *Inst. Phys. Conf. Ser.*, 96, 205 (1989).
168. J. P. Salerno, J. W. Lee, R. E. McCullough, and R. P. Gale, Epitaxial growth of GaAs on 4-inch diameter silicon substrates by OMCVD, *Mater. Res. Soc. Proc.*, 91, 119 (1987).

169. S. Sakai, New method to relax thermal stress in GaAs grown on Si substrates, *Appl. Phys. Lett.*, 51, 1069 (1987).
170. M. Yamaguchi, M. Tachikawa, M. Sugo, S. Kondo, and Y. Itoh, Analysis for dislocation density reduction in selective area grown GaAs films on Si substrates, *Appl. Phys. Lett.*, 56, 27 (1990).
171. A. A. Griffith, The phenomenon of rupture and flow in solids, *Phil. Trans. R. Soc.*, 221, 163 (1921).
172. A. Kelly, *Strong Solids* (Oxford: Clarendon Press, 1966), p. 45.
173. J. W. Matthews and E. Klockholm, Fracture of brittle films under the influence of misfit stress, *Mater. Res. Bull.*, 7, 213 (1972).
174. F. Semond, P. Lorenzini, N. Grandjean, and J. Massies, High-electron-mobility AlGaN/GaN heterostructures grown on Si(111) by molecular-beam epitaxy, *Appl. Phys. Lett.*, 78, 335 (2000).
175. H. M. Liaw, R. Venugopal, J. Wan, R. Doyle, P. Fejes, M. J. Loboda, and M. R. Melloch, Crack-free, single-crystal GaN grown on 100 mm diameter silicon, *Mater. Sci. Forum*, 338–342, 1463 (2000).
176. E. Feltin, B. Beaumont, M. Laügt, P. de Mierry, P. Vennéguès, H. Lahrèche, M. Leroux, and P. Gibart, Stress control in GaN grown on silicon (111) by metalorganic vapor phase epitaxy, *Appl. Phys. Lett.*, 79, 3230 (2001).
177. J. P. Zhang, H. M. Wang, M. E. Gaevski, C. Q. Chen, Q. Fareed, J. W. Yang, G. Simin, and M. Asif Khan, Crack-free thick AlGaN grown on sapphire using AlN/AlGaN superlattices for strain management, *Appl. Phys. Lett.*, 80, 3542 (2002).

6

Mismatched Heteroepitaxial Growth and Strain Relaxation: II. Graded Layers and Multilayered Structures

6.1 Introduction

Practical semiconductor devices often incorporate compositional grading or multilayer structures. These complex structures are designed to achieve the desired control of the strain and dislocation densities, as well as the necessary electronic or optoelectronic properties of the devices.

Many aspects of these device structures are similar to the case of a single uniform mismatched layer. For example, the growth may initially be pseudomorphic with lattice relaxation commencing when it becomes energetically favorable at a critical layer thickness. The lattice relaxation mechanisms, dislocation multiplication mechanisms, and dislocation interactions are similar to uniform layers. The lattice relaxation, if it occurs near equilibrium, is dictated by minimum energy considerations. In most cases, however, the lattice relaxation is controlled by the thermally activated glide of dislocations, and is considered to be limited by kinetics.

Differences between compositionally graded and uniform layers arise because in the graded layer the lattice mismatch and strain vary with distance from the interface. Whereas in the uniform layer the misfit dislocations are all located in the substrate interface with a linear density ρ_L, those in a graded layer are distributed throughout its thickness with an areal density ρ_A, and this areal density varies with position in the layer. The variation in lattice mismatch also gives rise to nonuniform strain throughout the layer.

This chapter develops the general theoretical framework for graded structures, including the critical layer thickness, equilibrium strain and misfit dislocation density, kinetically limited strain relaxation, and threading dislocation density behavior. These developments are quite general, and applicable to graded and multilayer structures. Then some specific graded and multilayer structures of interest are considered, including linearly graded, sublinearly graded, superlinearly graded, S-graded layers, and strained layer superlattices.

6.2 Critical Layer Thickness: General Case

Lattice relaxation can occur in a graded layer by the introduction of misfit dislocations if either the layer as a whole or any portion of it becomes unstable with respect to this process. Either way, we can consider that lattice relaxation commences once a critical

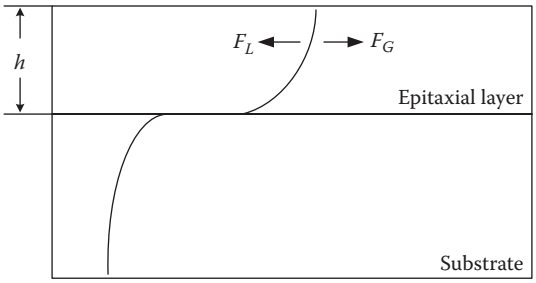


FIGURE 6.1
Forces acting on a grown-in dislocation in a compositionally graded layer.

layer thickness has been exceeded. As in the case of the uniform layer, force balance or minimum energy may be applied to determine the condition for lattice relaxation, or the critical layer thickness. An intriguing difference is that the graded layer may exhibit more than one critical layer thickness, as will be shown in the later sections.

We can determine the approximate critical layer thickness for such a layer using a simple force balance model, as illustrated in Figure 6.1. Following the approach presented by Matthews and Blakeslee[1] for a structure containing uniform strained layers, we assume that a grown-in dislocation will elongate at the interface to create a length of misfit dislocation once the glide force F_G exceeds the dislocation line tension F_L. Suppose that the lattice mismatch f at a distance z from the interface is

$$f(z) \equiv \frac{a_s - a_e(z)}{a_e(z)}, \tag{6.1}$$

where a_s is the lattice constant of the substrate and is the lattice constant of the epitaxial layer at a distance from the interface. The glide force acting on a grown-in dislocation in the coherently strained layer on a (001) diamond or zinc blende substrate with nonuniform strain $\varepsilon_{\parallel}(z) = f(z)$ is

$$F_G = \frac{2Gb(1+\nu)\cos\lambda}{(1-\nu)} \int_0^h f(z)dz, \tag{6.2}$$

where:

G is the shear modulus, $G = (C_{11} - C_{12})/2$
C_{11} and C_{12} are the elastic stiffness constants
b is the length of the Burgers vector
ν is the Poisson ratio, $\nu = C_{12}/(C_{11} + C_{12})$, for the (001) orientation
λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface
h is the layer thickness
z is the distance from the substrate interface

Although the material parameters b, C_{11}, and C_{12} and all vary with composition—and therefore distance from the interface—average values may be used in an approximate treatment.

If the grown-in dislocation glides to create a length of misfit dislocation at the substrate interface ($z = 0$), the line tension in this misfit segment that opposes the glide is

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{6.3}$$

where α is the angle between the Burgers vector and the line vector for the dislocations. Equating the glide and line tension forces at the critical layer thickness h_c, we obtain

$$\int_0^{h_c} f(z)dz = \frac{b\left(1-\nu\cos^2\alpha\right)\left[\ln\left(h_c/b\right)+1\right]}{8\pi(1+\nu)\cos\lambda}, \tag{6.4}$$

which can be solved numerically to determine the approximate critical layer thickness. It is important to point out that this derivation is based on the assumption that misfit dislocations are first introduced at the substrate interface. Although this is a good assumption for a uniform strained layer, it will be subject to some degree of error if the first misfit dislocations are introduced away from the interface, as in the case of a linearly graded[2] or exponentially graded[3] layer. Moreover, this simple model involves the use of effective values for parameters b, C_{11}, and C_{12}, even though these vary with composition in a graded layer.

In some graded layers, the initial misfit dislocations may be formed at a distance from the interface z_c, as shown in Figure 6.2. Then, the glide force acting on the dislocation at the critical layer thickness is modified to

$$F'_G = \frac{2Gb(1+\nu)\cos\lambda}{(1-\nu)}\int_{z_c}^{h} f(z)dz, \tag{6.5}$$

where z_c is distance from the interface where misfit dislocations initially form, and the line tension on an initial misfit dislocation is

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h-z_c}{b}\right)+1\right]. \tag{6.6}$$

Equating the glide and line tension forces at the critical layer thickness h_c, we obtain

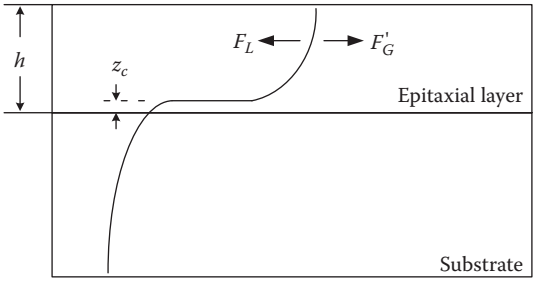


FIGURE 6.2
Formation of initial misfit dislocations at a finite distance from the interface in a compositionally graded layer.

$$\int_{z_c}^{h_c} f(z)dz = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi\left(1+\nu\right)\cos\lambda}\left[\ln\left(\frac{h_c - z_c}{b}\right)+1\right]. \tag{6.7}$$

This can be solved numerically for the critical layer thickness.

6.3 Equilibrium Strain and Misfit Dislocations: General Case

Bertoli et al.[4] have described a numerical approach for the determination of the equilibrium configuration of strain and misfit dislocations in a generalized semiconductor heterostructure. This approach is based on the minimization of the sum of the dislocation line energy and strain energy, and represents a generalization of the Tersoff[5] model for linearly graded structures.

In the general case of a compositionally graded semiconductor heterostructure with misfit dislocations of cross-sectional density $\rho(z)$, the in-plane strain is relaxed to

$$\varepsilon_{\|}(z) = f(z) + b'\int_0^z \rho(\xi)d\xi\,, \tag{6.8}$$

and the strain energy per unit area is

$$E_e = \int_0^h Y\left[f(z) + b'\int_0^z \rho(\xi)d\xi\right]^2 dz\,. \tag{6.9}$$

The dislocation density ρ is always positive, but b' may be positive or negative, depending on the sign of the mismatch strain that is being relieved. For a uniform mismatched layer, sign(b') = −sign(f). The energy of dislocations per unit area, assuming two orthogonal networks with equal cross-sectional density, is

$$E_d = 2\int_0^h F_d(z)\rho(z)dz, \tag{6.10}$$

where

$$F_d(z) \approx \frac{G(z)b^2(z)\left[1-\nu(z)\cos^2\alpha\right]}{4\pi\left(1-\nu(z)\right)}\left[\ln\left(\frac{h-z}{b(z)}\right)+1\right], \tag{6.11}$$

and G is the shear modulus, G = (C_{11} − C_{12})/2. Here, the dislocation energy is determined using a mean-field approach, without including the dislocation–dislocation interactions, similar to the work of Tersoff. The equilibrium configuration may be determined by minimizing E_d + E_e, and this can be solved numerically using the computational approach summarized below.

In the implementation of the generalized semiconductor heterostructure energy minimization numerical tool by Bertoli et al.,[4] a structure with arbitrary compositional profile

is approximated by N constant composition layers, each having a thickness no more than h_{max}. (h_{max} is set to a value corresponding to the desired spatial resolution, typically 5 nm or less.) The lattice constant and elastic constants for the j^{th} layer are $a_e[j]$, $C_{11}[j]$, and $C_{12}[j]$, and these may be estimated using linear interpolation (Vegard's law) for semiconductor alloys such as $Si_{1-x}Ge_x$ or $In_xGa_{1-x}As$. The lattice mismatch strain (with respect to the substrate) in the j^{th} layer is

$$f[j] = \frac{a_s - a_e[j]}{a_e[j]}, \tag{6.12}$$

the biaxial modulus for the j^{th} layer is

$$Y[j] = C_{11}[j] + C_{12}[j] - \frac{2C_{12}^2[j]}{C_{11}[j]}, \tag{6.13}$$

and the Poisson ratio for the j^{th} layer is

$$\nu[j] = \frac{C_{12}[j]}{C_{11}[j] + C_{12}[j]}. \tag{6.14}$$

The strain energy per unit area is found using

$$E_e = \sum_{j=1}^{N} \varepsilon^2[j] Y[j] h[j], \tag{6.15}$$

and the dislocation line energy is calculated by

$$E_d = \sum_{j=1}^{N} K[j] \rho[j] h[j], \tag{6.16}$$

where

$$K[j] = \frac{G[j] b^2[j] \{1 - \nu[j] \cos^2 \alpha\}}{2\pi(1 - \nu[j])} \left[\ln\left(\frac{\sum_{k=1}^{N} h[k] - \sum_{i=1}^{j} h[i]}{b[j]} \right) + 1 \right]. \tag{6.17}$$

In general, the compositional profile may include a combination of graded and uniform sublayers. Any sublayer thicker than h_{max} is subdivided. The strain profile is arbitrarily initialized to the pseudomorphic profile: $\varepsilon[j] = f[j]$ for each layer j. The misfit dislocation density is arbitrarily initialized to zero for every layer in the structure.

The basic process of finding the minimum energy configuration is as follows. Starting with the first layer ($j = 1$), the dislocation density in each layer j is adjusted up by Δ ($\rho[j] = \rho[j] + \Delta$). Accounting for this adjustment, the modified strain profile is calculated by

$$\varepsilon[n] = f[n] + \sum_{i=1}^{n} h[i]b'[i]\rho[i], \tag{6.18}$$

and the energy per unit area $E^{\uparrow}$ is calculated using

$$E = \sum_{n=1}^{N} \left\{ \varepsilon^2[n] Y[n] h[n] + K[n]\rho[n]h[n] \right\}. \tag{6.19}$$

Next, the dislocation density in layer j is adjusted back to its starting point ($\rho[j] = \rho[j] - \Delta$), the resulting strain profile is calculated, and the energy E^{-} is found. Finally, the dislocation density in the first layer is adjusted down by Δ ($\rho[j] = \rho[j] - \Delta$), the revised strain profile is found, and the energy $E^{\downarrow}$ is calculated. At this point, the values $E^{\uparrow}$, E^{-}, and $E^{\downarrow}$ are compared. The configuration corresponding to the lowest energy is adopted, and the process is repeated with the next layer, $j + 1$.

Once the dislocation density has been adjusted in each layer of the structure, the adjustment parameter Δ is reduced by a convenient multiplier (less than 1, and typically 0.95) and the process is repeated for each of the layers in the structure. This is continued until Δ has been reduced to a value below the desired dislocation density resolution Δ_{res}. (In the work of Bertoli et al., the value $\Delta_{res} = 50\ \text{cm}^{-2}$ was adopted.) The value of Δ is initialized to Δ_{start}, where

$$\Delta_{start} = \frac{\sum_{n=1}^{N} f[n]}{Nb'[1]h[1]}, \tag{6.20}$$

and the final solution is unaffected by this initial value as long as it is sufficiently large to allow convergence.

In some work, it was found that the constraint of adjusting the dislocation density in one layer at a time gave rise to nonphysical oscillations in the dislocation density profile, and to avoid this problem, it is necessary to adopt a *look-ahead energy minimizer*, which considers two or more layers at a time. In the case for which two layers are adjusted together, nine values of energy need to be compared: $E^{\uparrow\uparrow}$, $E^{\uparrow-}$, $E^{\uparrow\downarrow}$, $E^{-\uparrow}$, E^{--}, $E^{-\downarrow}$, $E^{\downarrow\uparrow}$, $E^{\downarrow-}$, and $E^{\downarrow\downarrow}$. Here, $E^{\uparrow\uparrow}$ is the energy corresponding to the case in which the dislocation density has been increased by Δ in both layer j and layer $j + 1$; $E^{\uparrow-}$ is the energy corresponding to the case in which the dislocation density has been increased by Δ in layer j but kept unchanged in layer $j + 1$, and so on. The configuration with the lowest energy is adopted out of these nine cases, and then the process is repeated with layers $j + 1$ and $j + 2$. A look-ahead energy minimizer could also consider three or more layers at a time. For example, consideration of three layers at a time involves comparing the energies for 27 permutations at each step. This greatly increases the computational time without any apparent improvement in the accuracy of the final solution, so a two-layer look-ahead energy minimizer was used in the original work by Bertoli et al.[4]

6.4 Kinetically Limited Strain Relaxation: General Case

Most heteroepitaxial device structures are not grown in equilibrium. Therefore, nonequilibrium models for dislocation flow must be applied to understand the kinetically

limited strain relaxation and misfit dislocation density. Threading dislocations, which are nonequilibrium defects, can only be understood using such a model.

Kujofsa et al.[6] extended the dislocation flow models previously developed for uniform layers to the general case of a nonuniform structure, including graded and multilayered structures. In that work, it was assumed that (1) the rate of lattice relaxation at a distance z from the interface is governed by the glide of misfit dislocations that are between the interface and the point z, and (2) the glide of dislocations at a distance z from the interface is governed by the glide force acting on threading arms of dislocations between the point z and the surface. Based on these starting assumptions, the effective stress at a distance z from the interface is

$$\tau_{\text{eff}}(z) = \left[\frac{2\cos\phi\cos\lambda}{h-z}\right]\int_z^h \left\{\frac{G(1+\nu)\left[\varepsilon_{\parallel}(\zeta)-\varepsilon_{\text{eq}}(\zeta)\right]}{(1-\nu)}\right\} d\zeta, \tag{6.21}$$

where:

- ϕ is the angle between the surface normal and the slip plane
- λ is the angle between the Burgers vector and the direction in the interface that is perpendicular to the intersection of the glide plane and the interface
- G is the shear modulus
- ν is the Poisson ratio
- $\varepsilon_{\parallel}$ is the in-plane strain
- ε_{eq} is the equilibrium in-plane strain
- h is the layer thickness

The equilibrium strain profile for an arbitrary heteroepitaxial structure may be determined by the energy minimization approach described in Section 6.3. Using the Alexander and Haasen model[7] for glide velocity and the Dodson and Tsao model[8] for dislocation multiplication, the rate of lattice relaxation at a distance z from the interface is given by

$$\frac{d\gamma(z)}{dt} = KBb\sin\alpha\sin\phi\tau_{\text{eff}}^2(z)\exp\left(-\frac{U}{k_BT}\right)\int_0^z\left[\rho_A(\zeta)+\rho_0\right]d\zeta, \tag{6.22}$$

where:

- K and B are constants
- b is the length of the Burgers vector
- U is the activation energy for dislocation glide
- k_B is the Boltzmann constant
- T is the temperature
- B is a constant where ρ_0 accounts for initial sources of misfit dislocations in a manner analogous to the γ_0 term used by Dodson and Tsao, and the areal density of misfit dislocations is given by

$$\rho_A(z) = \left(\frac{1}{b\sin\alpha\sin\phi}\right)\left(\frac{d\varepsilon_{\parallel}(z)}{dz}-\frac{df(z)}{dz}\right), \tag{6.23}$$

where $f(z)$ is the lattice mismatch, defined here as $f(z) \equiv (a_s - a(z))/a(z)$, where $a(z)$ is the relaxed lattice constant of the epitaxial layer at a distance z from the interface and a_s is the relaxed lattice constant of the substrate.

In a general heteroepitaxial structure, it is necessary to allow for the possibilities of varying growth temperature and growth rate, growth interruptions or annealing steps, and nonuniform composition. All of these effects can be treated by approximating the structure by a stack of sublayers, each of which is assumed to have uniform properties, in much the same way that equilibrium modeling has been approached. An advantage of this approach is that it can account for the time evolution of the equilibrium strain profile, thus enabling a more accurate description of uniform layers, as well as compositionally graded structures and multlilayers.

To implement such a model, the structure is considered to grow in a step-by-step fashion, one sublayer per step. Cheruku et al.[9] described an algorithm for an implementation of this type. At step 1 in the plastic flow modeling process, the structure is considered to comprise only the substrate and sublayer 1. The layer is assumed to be deposited coherently (pseudomorphically) with its in-plane strain equal to the lattice mismatch: $\varepsilon_{\|}[1] = f[1]$. The equilibrium strain is calculated by an energy minimization procedure as described in Section 6.3. If the equilibrium strain differs from the coherency strain, this results in a nonzero effective stress:

$$\tau_{\text{eff}}[1] = \frac{2G(1+\nu)\cos\phi\cos\lambda}{(1-\nu)}\left\{\varepsilon_{\|}[1] - \varepsilon_{\text{eq}}[1]\right\}. \tag{6.24}$$

Layer 1 is considered to anneal for a duration equal to its growth time, $t[1] = h[1]/g[1]$, during which the lattice relaxation is

$$\Delta\gamma[1] = \text{sign}\left\{\varepsilon_{\text{eq}}[1] - \varepsilon_{\|}[1]\right\} KBb \sin\alpha \sin\phi \tau_{\text{eff}}^2[1] \exp\left(-\frac{U}{k_B T[1]}\right) \rho_0 h[1] t[1]. \tag{6.25}$$

The in-plane strain $\varepsilon_{\|}[1]$ is replaced by $\varepsilon_{\|}[1] + \Delta\gamma[1]$, the areal misfit dislocation density in sublayer 1 is set to $\rho[1] = (\varepsilon_{\|}[1] - f[1])h[1]/(b \sin\alpha \sin\phi)$, and the process proceeds with step 2.

At any step m (for which $m > 1$), the m^{th} sublayer is added to the structure and is assumed to be strained in such a way that it is coherent with the $(m-1)^{\text{th}}$ layer. The starting value of strain in the m^{th} layer is therefore $\varepsilon_{\|}[m] = \varepsilon_{\|}[m-1] + f[m] - f[m-1]$. To account for any possible temperature change at the m^{th} layer, the in-plane strain in each of the layers is adjusted by an amount $\Delta\varepsilon_{Th}[i]$ given by

$$\Delta\varepsilon_{\text{Th}}[i] = \int_{T[m-1]}^{T[m]} \left(\alpha_S - \alpha[i]\right) dT, \tag{6.26}$$

where $T[m-1]$ and $T[m]$ are the growth temperatures for the $(m-1)^{\text{th}}$ and m^{th} sublayers, respectively, α_s is the thermal coefficient of expansion for the substrate, and $\alpha[i]$ is the thermal coefficient of expansion for the i^{th} sublayer. Next, the effective stress is found for each of the sublayers; in the n^{th} layer during the m^{th} step of the relaxation process, it is

$$\tau_{\text{eff}}[n] = \frac{2\cos\phi\cos\lambda \sum_{i=n}^{m} \left\{G[n]h[i](1+\nu[n])\left(\varepsilon_{\|}[n] - \varepsilon_{\text{eq}}[n]\right)/(1-\nu[n])\right\}}{\sum_{i=n}^{m} h[i]}. \tag{6.27}$$

Using this effective stress profile, the lattice relaxation is calculated for each of the sublayers up to sublayer m. For the n^{th} layer during the m^{th} step of the relaxation process, the lattice relaxation is

$$\Delta\gamma[n] = \text{sign}\left\{\varepsilon_{\text{eq}}[1] - \varepsilon_{\parallel}[1]\right\} KBb \sin\alpha \sin\phi \tau_{\text{eff}}^2[n] \exp\left(-\frac{U}{k_B T[n]}\right) t[n] \sum_{i=1}^{n} \left[\left(\rho[i] + \rho_0\right) h[i]\right]. \quad (6.28)$$

Once $\Delta\gamma[1]$, $\Delta\gamma[2]$, …, $\Delta\gamma[m]$ have been calculated, the strain profile is adjusted by replacing each in-plane strain $\varepsilon_{\parallel}[n]$ with $\varepsilon_{\parallel}[n] + \Delta\gamma[n]$. Finally, the misfit dislocation density profile is adjusted according to

$$\rho[i] = \left\{ \begin{array}{ll} \dfrac{\left(\varepsilon_{\parallel}[i] - f[i]\right)}{bh[i] \sin\alpha \sin\phi}, & i = 1; \\ \dfrac{\left(\varepsilon_{\parallel}[i] - \varepsilon_{\parallel}[i-1] - f[i] - f[i-1]\right)}{bh[i] \sin\alpha \sin\phi} & i \neq 1. \end{array} \right\}, \quad (6.29)$$

and the process proceeds to step $m + 1$.

A growth rest is modeled as a "sublayer" having an associated time and temperature but zero thickness. (In practice, a layer of negligible thickness is used.) A linear temperature ramp can be treated approximately in the same way, using the approach adopted for the treatment of temperature ramping in diffusion furnaces.[10] If the temperature is ramped from T_1 to T_2 with a ramp time t_{ramp}, the lattice relaxation can be approximated by using an annealing step at T_1 for an effective time t_{eff} given by

$$t_{\text{eff}} = \frac{k_B\left(T_2^2 - T_1^2\right)}{U\left(T_2 - T_1\right)} t_{\text{ramp}}, \quad (6.30)$$

where:

k_B is the Boltzmann constant
u is the activation energy for dislocation glide

6.5 Threading Dislocation Densities: General Case

Models for threading dislocation densities in step-graded and linearly graded layers will be considered in the following sections. In the general case, a plastic flow model can be applied to a general laminar structure for the determination of the threading dislocation density profile. Kujofsa et al. developed a general numerical model for threading dislocations in compositionally graded structures including two important misfit–threading dislocation interactions: (1) the introduction of dislocation half loops[11] and (2) the bending over of existing threading dislocations at mismatched interfaces.[1] When misfit dislocations are created by the introduction of half loops, each misfit dislocation segment of length L_{MD} is associated with two threading segments that intersect the surface, thereby adding to the threading dislocation population. In contrast, the creation of misfit segments

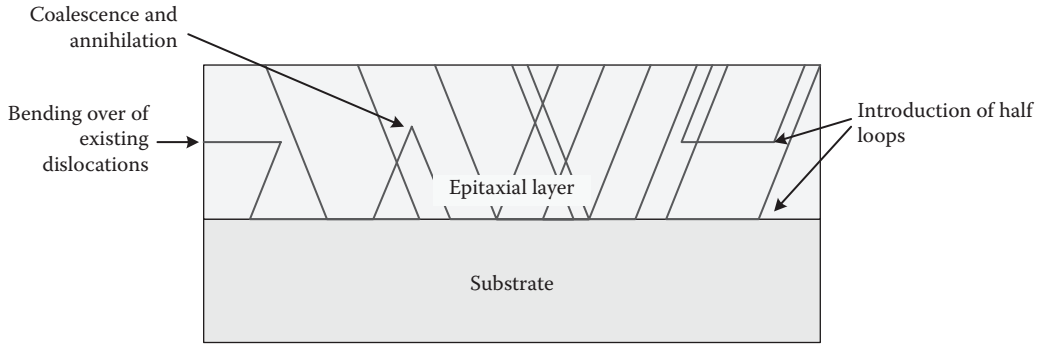


FIGURE 6.3
Threading dislocation interactions in a mismatched heteroepitaxial layer. (Reprinted from T. Kujofsa et al., *J. Electron. Mater.*, 41, 2993 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

by the bending over of grown-in threading dislocations takes away from the number of threading dislocations, because dislocations bent over at an interface can combine in an annihilation reaction or reach the edge of the sample. In addition to misfit dislocation–mediated annihilation, there can be first-order and second-order coalescence and annihilation reactions involving only threading dislocations, as modeled by Tachikawa and Yamaguchi[12] and Romanov et al.[13] and described in Chapter 5. The three basic types of dislocation interactions are shown schematically in Figure 6.3, and the resulting threading density may be modeled by

$$\frac{dD(z)}{dz} = \frac{4\rho_A(z)}{L_{MD}(z)\text{sign}\int_0^z \rho_A(\zeta)d\zeta} - C_1 D(z) - C_2 D^2(z), \tag{6.31}$$

where:

$\rho_A(z)$	is the density of misfit dislocations
$L_{MD}(z)$	is the length of misfit dislocation segments
z	is the distance from the interface

The first term accounts for the interactions between misfit and threading dislocations. Considering mechanism 1, new misfit dislocations are introduced via half loops if the new misfit dislocations have the same sense (relax the same sign of lattice mismatch) as the underlying misfit segments. This corresponds to the case of $\text{sign}(\rho_A(z)) = -sign\int_0^z \rho_A(\zeta)d\zeta$ and results in positive $dD(z)/dz$. With respect to mechanism 2, misfit dislocations are produced by the bending of existing threading dislocations if these misfit dislocations have the opposite sense (relax the opposite sign of mismatch) compared with the underlying misfit segments. This corresponds to $\text{sign}(\rho_A(z)) = -sign\int_0^z \rho_A(\zeta)d\zeta$ and results in negative $dD(z)/dz$. The second term accounts for first-order reactions and the third term accounts for second-order annihilation–coalescence reactions involving threading dislocations. In the work of Kujofsa et al.,[6] the values $C_1 = 200\ \text{cm}^{-1}$ and $C_2 = 1.8 \times 10^{-5}$ cm were adopted, as given by Tachikawa and Yamaguchi[12] for GaAs/Si (001), based on the finding that ZnSe/GaAs (001) and GaAs/Si (001) heteroepitaxial layers exhibit approximately the same thickness dependence of the threading dislocation density.[14,15] Application of this model requires the use of a kinetic model for lattice relaxation, such as the one described in the

previous section, in order to predict the distribution of misfit dislocations and the lengths of the misfit segments.

6.6 Step-Graded Layer

In a step-graded buffer, the overall change in lattice constant is divided between several steps. This approach has two advantages relative to the uniform buffer: the individual steps in the lattice constant may be made small enough to preserve layer-by-layer growth, and the misfit dislocations are spread out in the growth direction by placing them at multiple abrupt interfaces. Step-graded buffers have been investigated in a number of material systems, including $In_xGa_{1-x}As$,[16–21] $In_xAl_{1-x}As$,[20,22–34] $In_xAl_yGa_{1-x-y}As$,[34–36] and $Al_xGa_{1-x}Sb_yAs_{1-y}$[37,38] on GaAs (001) substrates, $InAs_yP_{1-y}$[39–42] on InP (001), and $In_xAl_{1-x}Sb$[43] on GaSb (001).

There has been relatively little modeling work with step-graded buffers, and their theoretical behavior is less firmly established than for the linearly graded case, but some important properties are known regarding their dislocation and strain behavior. Abrahams et al.[44] considered a simple model for a step-graded structure and argued that the threading dislocation density would reach a steady-state value that depends on the average grading coefficient, similar to the case of linear grading. This model is illustrated schematically in Figure 6.4, in which misfit dislocations are created by the jogging of threading dislocations at each interface. If the average length of misfit segments is fixed, the steady-state threading dislocation density will be proportional to the amount of mismatch at each step (the average grading coefficient). Experimental results tend to bear this out, and show the absence of misfit dislocations above the top step interface. Experiments

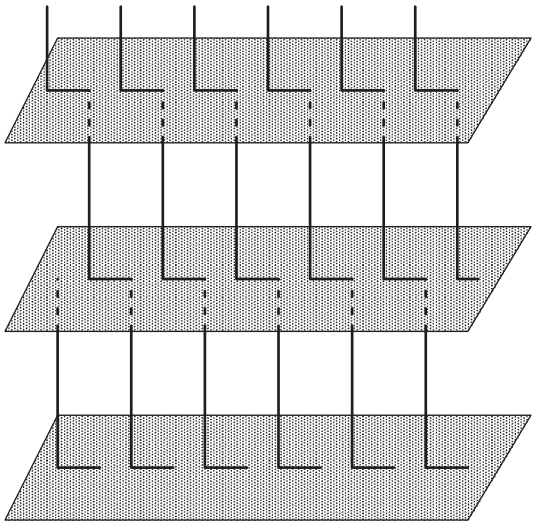


FIGURE 6.4
Abrahams model for the step-graded buffer, predicting a steady-state threading dislocation density proportional to the average grading coefficient. (Adapted from M. S. Abrahams et al., *J. Mater. Sci.*, 4, 223 [1969]. With permission. Copyright 1969, Springer.)

have also shown maximum residual strain in the topmost layer of a step-graded buffer, and this is to be expected if lattice relaxation proceeds with the deposition of each successive step.[17] Experiments with $In_xAl_{1-x}As$ on GaAs showed that step-graded and linearly graded buffers had similar values of residual surface strain.[31] Therefore, we can identify several similarities to linearly graded layers: (1) misfit dislocations are absent from the topmost portion of the buffer, in a misfit dislocation free zone (MDFZ); (2) there is a relatively large built-in strain in the top material; and (3) overshoot of the buffer layer composition (in other words, a reverse step of the device layer) should allow growth of a nearly strain-free device layer, if its relaxed lattice constant matches the in-plane lattice constant of the strained layer at the top of the buffer. An important difference between step-graded and linearly graded buffers is that the step-graded buffer contains an MDFZ in each step, whereas a linearly graded buffer contains at most two MDFZs. Experimental investigations of step-graded buffers have involved 10 or more steps[23,30,39] and as few as 2 steps.[21] If a large number of small steps are employed, the behavior of the step-graded layer will approach that of the linear buffer,[30] but otherwise, we should expect the abrupt interfaces to give rise to characteristics unique to step grading.

6.6.1 Lattice Relaxation and Residual Strain in a Step-Graded Layer

Considering force balance on a grown-in dislocation, one can see that the average residual strain is the same as that in a uniform layer of the same total thickness. The deposition of each successive step gives rise to lattice relaxation of the underlying layers[29] to maintain the inverse relationship between average strain and thickness, and therefore the residual strain is concentrated in the top layer of the step-graded buffer. For example, in a step-graded $InAs_yP_{1-y}$ structure grown on InP (001) by metalorganic vapor phase epitaxy (MOVPE) for the realization of a mid-infrared laser diode, Kirch et al.[39] found that the top layer (500 nm $InAs_{0.68}P_{0.32}$) contained 10% of the coherency strain—much greater than the equilibrium strain for a 500 nm constant composition layer. The built-in strain at the top of the step-graded buffer helps sweep out threading dislocations, generally leading to a reduced threading dislocation density compared with the case of a uniform buffer. To avoid new misfit dislocations in the device layer, the buffer can be designed with compositional overshoot so that the relaxed device layer matches the in-plane lattice constant of the buffer layer.[29,30,32,45] For $In_xGa_{1-x}As/In_yAl_{1-y}As$ quantum wells on GaAs substrates, it is possible to control the residual strain in the quantum wells by varying the amount of overshoot in the step-graded buffer.[30] Exact matching is difficult to achieve in practice, unless the relaxation dynamics can be predicted accurately, but could be facilitated by the use of *in situ* stress measurements.[29]

The residual strain and lattice relaxation in step-graded buffers are usually assessed by *ex situ* x-ray methods, including x-ray reciprocal space mappings (RSMs) and high-resolution x-ray rocking curves (HRXRCs). Dynamical simulations of HRXRCs allow depth profiling of composition and strain as long as symmetrical and asymmetrical reflections are used.[46–48] Recently, it has been shown that this method may be extended to include the profiling of dislocation densities.[49,50] In most cases, the dynamical simulations are made necessary by the lack of a one-to-one correspondence between x-ray diffraction peaks and lattice constants present in the sample.[46] In some step-graded layers, though, a separate peak may be resolved for each layer, and it becomes possible to extract approximate strain data directly from the peak positions. This is the case for the HRXRC of Figure 6.5, measured from an overshoot step-graded $In_xAl_{1-x}As$ structure grown on GaAs (001) by molecular beam epitaxy (MBE) and incorporating nine 100 nm steps with indium mole fractions of 0.05, 0.15,

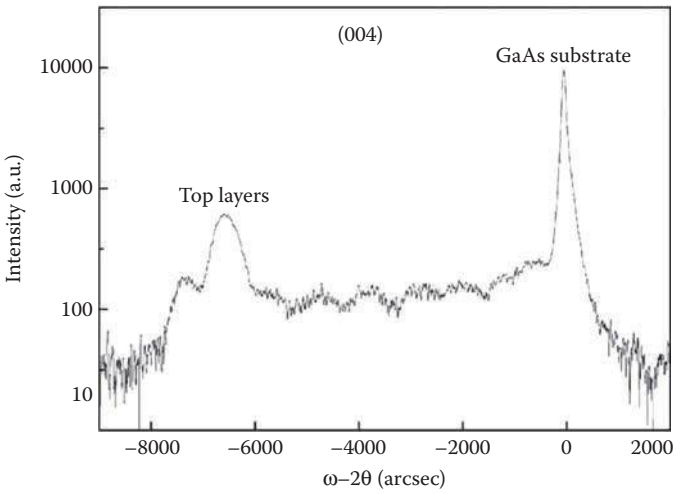


FIGURE 6.5
004 high-resolution x-ray diffraction (HRXRD) rocking curve for 1 μm $In_{0.75}Al_{0.25}As$ grown on a step-graded $In_xAl_{1-x}As$ buffer layer with an indium compositional overshoot of $\Delta x_{os} = 0.1$. (Reprinted from Z. Jiang et al., *Appl. Surf. Sci.*, 254, 5241 [2008]. With permission. Copyright 2008, Elsevier.)

0.25, 0.35, 0.45, 0.55, 0.65, 0.75, and 0.85. A 1 μm uniform layer of $In_{0.75}Al_{0.25}As$ was grown on top of the step-graded buffer. Analysis of the x-ray rocking curve, after cancelling out the effect of crystallographic tilting by use of the rocking curves measured at opposing azimuths, revealed that the top (device) layer was completely relaxed, within experimental error, as a consequence of the overshoot design (98% relaxation).[32]

The surface strain of a graded structure may be determined using micro-Raman spectroscopy. Song et al.[51] applied this method to characterize the strain in a GaN cap layer grown on a graded InGaN buffer, which was in turn grown on a GaN buffer on a sapphire substrate. The graded structure was grown by MOVPE, starting with a 25 nm low-temperature GaN nucleation layer, and then a 500 nm GaN buffer was grown at a conventional temperature of 1000°C. A step-graded InGaN buffer was next grown at 750°C, with five 200 nm InGaN layers, with the indium mole fraction increased in five equal steps to 5%. Thin 10 nm "insertion" layers of GaN were placed between the InGaN layers of the step-graded buffer, grown at 750°C and annealed at 1000°C. The GaN cap was 100 nm thick and grown at 1000°C. For comparison with the step-graded structure, a control sample with the same total thickness of GaN was also grown. The surface strain was characterized by measuring the Raman shifts of the GaN E_2(TO) phonon. The in-plane strain was calculated by

$$\varepsilon_{xx} = \frac{\Delta\omega_\lambda}{2p_\lambda + q_\lambda r}, \tag{6.32}$$

where:

$\Delta\omega_\lambda$ is the observed shift of the E_2(TO) phonon mode
p_λ $= -810 \pm 25$ cm^{-1} and $q_\lambda = -920 \pm 60$ cm^{-1} are the deformation potential constants
r $= -2C_{13}/C_{33}$ is the ratio of out-of-plane to in-plane strains for the case of biaxial strain

The out-of-plane strain can then be found from $\varepsilon_{zz} = r\varepsilon_{xx}$. Based on this analysis, Song et al.[51] found that the inclusion of the graded InGaN buffer affected partial strain compensation in the GaN cap, resulting in a lower compressive stress than that observed in the

control sample by ~3×10^{-4}. They applied a simple layer-by-layer analysis to calculate the strain in the i^{th} layer of the structure,

$$\frac{a_{i0} - a_i}{a_i} = \varepsilon_{Li} + \varepsilon_{Ti}, \tag{6.33}$$

where:

a_i and a_{i0} are the in-plane lattice constant and relaxed lattice constant
ε_{Li} and ε_{Ti} are the mismatch strain and thermal strain, respectively, of the i^{th} layer

The mismatch strain was assumed to be

$$\varepsilon_{Li} = \frac{a_{i-1} - a_{i0}}{a_{i0}}, \tag{6.34}$$

where a_{i-1} is the in-plane lattice constant in the $(i - 1)$th layer and the thermal strain was assumed to be

$$\varepsilon_{Ti} = \int_{T_G}^{T_R} (\alpha_{i-1} - \alpha_i)\, dT, \tag{6.35}$$

where:

α_{i-1} and α_i are the thermal expansion coefficients for the $(i - 1)^{th}$ and i^{th} layers
T_G and T_R are the growth and room temperatures

This simple model was able to predict the strain compensation of the graded layer without invoking kinetic considerations.

In situ stress measurements are also of great interest, because the *ex situ* x-ray measurements do not reveal the time evolution of the strain relaxation. One method is the multibeam optical stress sensor (MOSS), in which multibeam laser illumination of the growing surface allows determination of the sample curvature and, indirectly, the average stress in the growing film. Although depth profiling is not possible with the MOSS technique, it does allow determination of the average stress as a function of time (or thickness) while the structure is being grown, and this can provide valuable insight into the relaxation dynamics. Lynch et al.[29] applied MOSS to step-graded $In_xAl_{1-x}As$ buffers with 250 nm thick steps on GaAs substrates to show that the relaxation of underlying layers proceeds with the deposition of each successive step, and that each step quickly reaches a near-constant stress early in its growth. The latter finding implies that it might be possible to use much thinner steps in a step-graded buffer design. Figure 6.6 shows the compositional grading profile, the stress–thickness product as a function of thickness, and the average stress as a function of thickness for this $In_xAl_{1-x}As$ step-graded buffer.

Some studies have revealed nearly complete relaxation in step-graded buffers and device structures, even without the use of overshoot, as long as the device layer is sufficiently thick, but these structures may contain misfit dislocations in the device layers. Shang et al.[22] utilized x-ray RSM for a step-graded $In_xAl_{1-x}As$ structure (100 nm steps, $\Delta x = 0.1$) with a 1 μm uniform layer of $In_{0.52}Al_{0.48}As$ on top to show that all layers were ~98% relaxed without using overshoot. Alternatively, use of thick layers in the step-graded buffer results in a high degree of lattice relaxation at the top of the buffer, in some cases removing the need

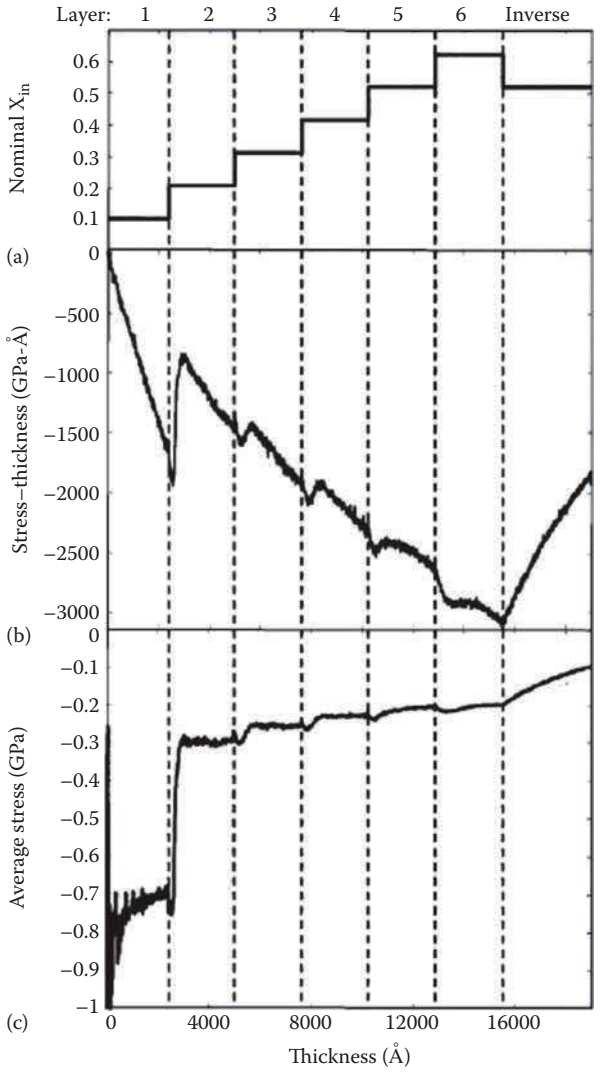


FIGURE 6.6
(a) Compositional grading profile, (b) stress–thickness product as a function of thickness, and (c) average stress as a function of thickness for an $In_xAl_{1-x}As$ step-graded structure. The stress–thickness product was determined by MOSS measurements. (Reprinted from C. Lynch et al., *J. Vac. Sci. Technol. B*, 22, 1539 [2004]. With permission. Copyright 2004, American Vacuum Society.)

for overshoot. Chen et al.[52] demonstrated 96% relaxation in the top layer of an $In_xGa_{1-x}As$ step-graded buffer grown on GaAs by MBE using relatively thick (0.3 μm) step layers and achieved excellent two dimensional electron gas (2DEG) mobility ($\mu_n = 35{,}000$ cm^2/Vs with $n_s = 1.21 \cdot 10^{12}$ cm^{-2} with $x = 0.3$ at 77K) without the need for overshoot.

6.6.2 Misfit and Threading Dislocations in a Step-Graded Layer

The body of reported work raises the question of whether step-graded or continuously graded buffer layers will be better for the reduction of the threading dislocation density

in metamorphic devices. Though there have been few direct comparisons, the answer to this question seems to depend on the material system and the ending composition. Some results suggest that step grading can be superior, confining dislocations to lower layers, if dislocations move more easily or have lower energy in material with those compositions. Arguments of this type have been made on the basis of yield strength and hardness. For step-graded $In_xGa_{1-x}As$ on GaAs (001), Krishnamoorthy et al.[18] showed that new dislocations were absent from the topmost layer as long as the change in composition was less than 0.18 per step, and they attributed this to the increasing yield strength with indium content. However, this approach was only effective up to an indium mole fraction of ~0.50, beyond which the yield strength actually decreases. Jeong et al.[23] made similar observations with step-graded $In_xAl_{1-x}As$ on GaAs (001). Though step grading was effective in reducing the threading dislocation density up to an alloy composition of ~0.50, they found that step-graded layers with higher indium concentrations yielded threading dislocation densities similar to those for uniform buffers of the same total thickness. They attributed this result to alloy hardening, which results in maximum hardness in the middle of the compositional range.

Based on the results above, it is interesting to consider the behavior of step-graded layers of $In_xGa_{1-x}As$ or $In_xAl_{1-x}As$ with high indium content, which have been investigated by Gozu et al.,[53] Mendach et al.,[27] Heyn et al.,[28] and Jiang et al.[32] Gozu et al.[53] fabricated an $In_{0.75}Ga_{0.25}As/In_{0.75}Al_{0.25}As$ high-electron-mobility transistor (HEMT) on GaAs using a step-graded $In_xAl_{1-x}As$ buffer, and although no threading dislocation densities were given, the device exhibited very high mobility, (μ_n = 397,000 cm^2/Vs with $n_s = 1.0 \cdot 10^{12}$ cm^{-2} at 4.2 K. Mendach et al.[27] grew an $In_{0.75}Ga_{0.25}As/In_{0.75}Al_{0.25}As$ quantum well on a step-graded $In_xAl_{1-x}As$ buffer on GaAs (001) having 100 nm steps with indium mole fractions of 0.15, 0.25, 0.35, 0.45, 0.55, and 0.65. The surface threading dislocation density was measured by plan-view transmission electron microscopy (PVTEM) to be 3.6×10^9 cm^{-2}. In an otherwise similar structure containing an SLS under the step-graded buffer, a lower threading dislocation density of 1.0×10^9 cm^{-2} was obtained, and under illumination, the 2 K Hall mobility was (μ_n = 400,000 cm^2/Vs with $n_s = 5.6 \cdot 10^{11}$ cm^{-2}. Heyn et al.[28] used this same step-graded structure with the inserted strained layer superlattice to fabricate a shallow-channel InAs HEMT, with the channel 16.5 nm below the surface and exhibiting 4 K channel mobility under illumination of (μ_n = 160,000 cm^2/Vs with $n_s = 6.3 \cdot 10^{11}$ cm^{-2} at 4.2 K. Figure 6.7 shows XTEM images of these structures (a) with and (b) without the SLS, showing the difference in threading dislocation behavior. Jiang et al. grew a uniform 1 μm $In_{0.75}Al_{0.25}As$ layer on a step-graded buffer incorporating overshoot with a maximum indium concentration of 0.85.[32] XTEM revealed that most threading dislocations were confined to the step-graded buffer, but the surface dislocation density was still a relatively high 2×10^8 cm^{-2}. These threading dislocation densities are similar to those that would be expected with uniform buffers having the same total thickness, suggesting weak dislocation reduction behavior in these high-indium materials, even though high electron mobilities have been obtained at cryogenic temperatures using them.

Quaternary step-graded buffers have also been investigated, such as $Al_xGa_{1-x}As_{1-y}Sb_y$ on GaAs for the realization of InAs QD lasers.[38] Using eight steps having fixed aluminum (x = 0.5) and a maximum antimony mole fraction y = 0.24, Xin et al. showed by XTEM that dislocations were absent from the partially relaxed top layer. Balakrishnan et al.[37] demonstrated InAs quantum dashes with room temperature emission at ~2.0 μm on a step-graded $Al_{0.5}Ga_{0.5}As_{1-y}Sb_y$ buffer with 16 steps and a top composition of y = 0.46. It was reported that this structure had a low threading dislocation density of $<10^5$ cm^{-2} at the top of the step-graded buffer as determined by TEM.

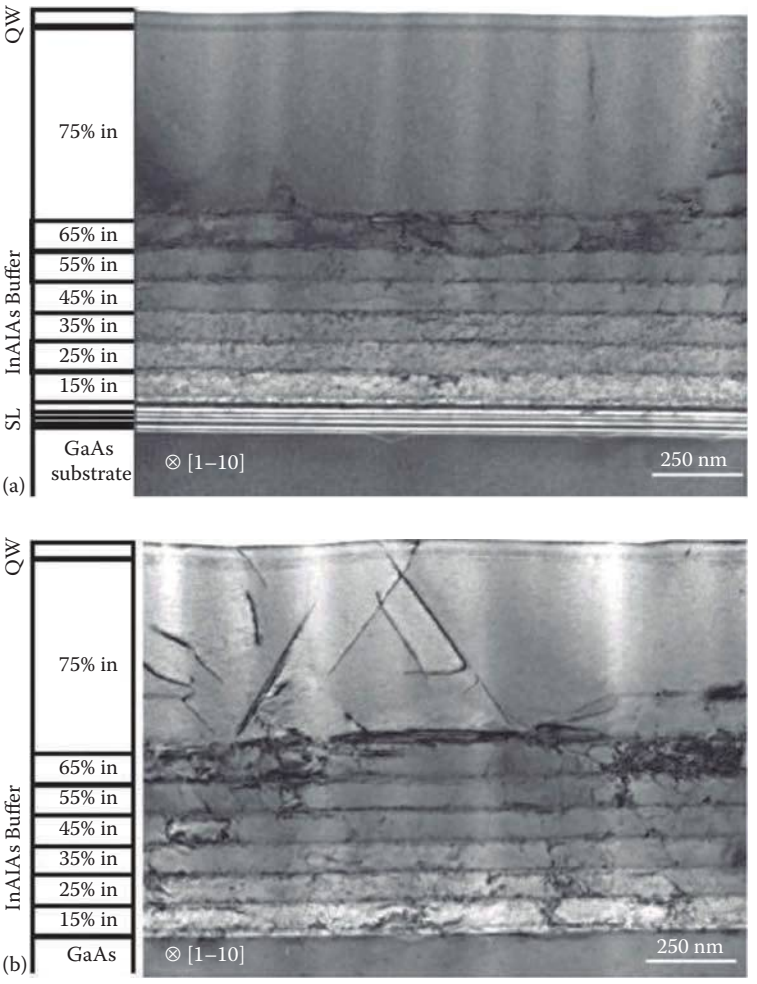


FIGURE 6.7
XTEM images of step-graded $In_xAl_{1-x}As$ structures grown on GaAs (001) substrates by MBE (a) with an inserted strained layer superlattice and (b) without a strained layer superlattice. (Reprinted from Ch. Heyn et al., *J. Cryst. Growth*, 251, 832 [2003]. With permission. Copyright 2003, Elsevier.)

The steady-state threading dislocation density is reached in the early stages of growth, so the conditions of growth at the first interface are critical in determining the surface dislocation density. Sometimes, the first step in composition is made smaller than the rest, or grown at a lower temperature, to take advantage of this fact. It has also been clearly shown that the insertion of a strained layer superlattice below a step-graded buffer establishes longer misfit segments, enhances lattice relaxation in the buffer, and reduces the surface dislocation density.[27,28,33]

Several studies have shown that low-temperature growth helps to confine threading dislocations to the metamorphic buffer. Grider et al.[54] showed that, for $In_xGa_{1-x}As$ HEMTs ($0.2 \leq x \leq 0.5$) grown on GaAs substrates by MBE with a step-graded $In_xGa_{1-x}As$ buffer, low-temperature buffer growth at 300°C resulted in a structure in which threading dislocations were confined to the buffer and the 77K electron mobility was >22,000 cm^2/Vs at $x = 0.4$. When the buffer was grown at the same temperature as the device, 500°C, there

was a high density of threading dislocations propagating through the device, as seen by PVTEM, and the 77 K mobility was lower by about a factor of three.

In some studies, high-temperature annealing has been applied between layers of the step-graded buffer to control threading dislocation propagation and reduce the surface density. Huo et al.[55] used *in situ* annealing at 500°C–540°C between layers of an $In_xGa_{1-x}As$ step-graded buffer grown on GaAs by MBE at 400°C. After growth of four 200 nm steps, with a top indium mole fraction of 0.4, a coherently strained 10 nm layer of Ge was grown on top, and no threading dislocations were observed in the Ge by XTEM, indicating that the surface density of dislocations was less than 10^8 cm^{-2}. Mi et al.[56] inserted 1.5 nm AlAs layers between the steps of a low-temperature (390°C) step-graded $In_xGa_{1-x}As$ buffer grown on a GaAs substrate by MBE to allow high-temperature (700°C) annealing of the $In_xGa_{1-x}As$ layers. This resulted in a sufficiently low threading dislocation density for the fabrication of a quantum dot laser diode emitting at 1.52 μm.

Some evidence indicates that lower threading dislocation densities may be obtained by reducing the average grading coefficient, either by using small compositional steps or thicker step layers. Typical step thicknesses are 100 nm, and typical average grading coefficients are 500–1000 cm^{-1}. As an example, a step-graded $In_xAl_{1-x}As$ buffer with 100 nm steps and an increment in the indium composition of $\Delta x = 0.1$ has an average grading coefficient of ~700 cm^{-1}. Czaban et al.[41] found, for uniform $InAs_yP_{1-y}$ material grown on an InP substrate with a step-graded $InAs_xP_{1-x}$ buffer, that doubling the step thickness (from 50 to 100 nm) increased the room temperature photoluminescence (PL) intensity from the device layers by 40%. He et al.[57] reported an extremely low etch pit density (EPD) of 5.0×10^3 cm^{-2}, comparable with the GaAs substrates, for a laser structure on a step-graded $In_xGa_{1-x}As$ buffer using thick steps (200 nm) and a small increment in the indium mole fraction ($\Delta x = 0.02$), resulting in an average grading coefficient of ~70 cm^{-1}.

6.6.3 Morphology and Surface Roughening in a Step-Graded Layer

Step-graded device structures typically exhibit cross-hatch morphology with corrugations along the [110] and [1$\bar{1}$0] directions,[54] as shown in the optical micrograph of Figure 6.8 for a mid-infrared laser diode on an InP(001) substrate with a step-graded $InAs_yP_{1-y}$ buffer

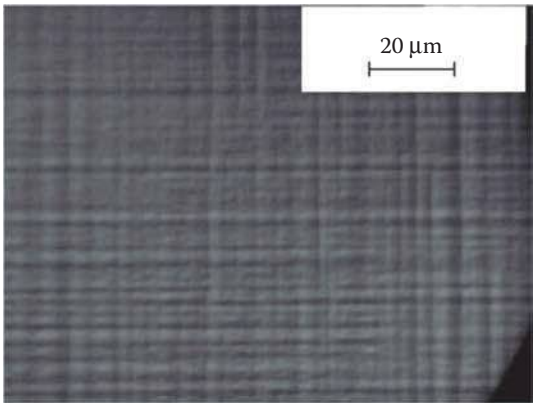


FIGURE 6.8
Optical micrograph showing cross-hatch pattern on the surface of a 500 nm InAsP/200 nm InGaAs SCH/InAs quantum well laser diode grown on a step-graded $InAs_yP_{1-y}$ buffer layer on an InP (001) substrate. (Reprinted from J. Kirch et al., *J. Cryst. Growth*, 312, 1165 [2010]. With permission. Copyright 2010, Elsevier.)

layer.[39] It has been shown that the average period of surface undulations corresponds approximately to the separation of misfit dislocations, projected into the interface and observed by PVTEM.[30] Often, the undulations are asymmetric and indicate asymmetric misfit dislocation densities. In addition to the normal cross-hatch, deep grooves have been found parallel to the [1$\bar{1}$0] direction in structures containing tensile residual strain in the device layers.[30] The cross-hatch arises from the influence of the misfit dislocation strain fields at the growing surface, which can induce compositional or growth rate variations. In a uniform layer, the misfit dislocations are buried at the interface, but the cross-hatch may persist in layers much thicker than the effective range of the misfit strain fields, due to secondary strain fields associated with height and compositional variations. In a step-graded layer, new misfit dislocations are introduced at each step, so the cross-hatch generally becomes more pronounced as the layer is grown. Increasing the separation of misfit dislocations also increases the effective range of their strain fields, which gives rise to cross-hatch. For these reasons, the associated surface roughness tends to be more problematic in graded buffers than in uniform composition layers.

The surface roughness associated with cross-hatch may be reduced by low-temperature growth,[58] high-temperature annealing,[55] or the use of surfactants, as has been shown qualitatively in Nomarski images and quantitatively by atomic force microscopy (AFM). The underlying physics involves manipulation of the surface mobility for adatoms. For step-graded $In_xAl_{1-x}As$ grown on GaAs by MBE with a sublinear profile and top indium composition of 0.35, studied by Nomarski microscopy, Shen et al.[58] found cross-hatch along both ⟨110⟩ directions for growth at 500°C, corrugations along the [1$\bar{1}$0] direction for growth at 450°C and 350°C, but an absence of corrugations and cross-hatch for low-temperature growth at 250°C. For sublinear step-graded $In_xAl_{1-x}As$ grown on GaAs (001) by MBE for the fabrication of $In_xGa_{1-x}As/In_yAl_{1-y}As$ HEMTs, Mishima et al. found that reduced temperature growth of step-graded $In_xAl_{1-x}As$ at 350°C resulted in superior morphology compared with deposition at 400°C.[25] For step-graded $In_xGa_{1-x}As$ grown on GaAs by MBE at 400°C and annealed between steps at 500°C–540°C, Huo et al.[55] found that the root mean square (rms) surface roughness in a 10 nm Ge top layer increased with the number of steps and top indium content, but was only 1.01 nm using four steps and a maximum indium composition of 0.4.

Three-dimensional growth is another cause of surface roughening, but lowering the growth temperature has been reported to extend the compositional regime for two-dimensional growth of some mismatched ternary alloys. This behavior has been reported in the case of a single $In_{0.1}Al_{0.9}As$ layer on GaAs by MBE, which could be grown in a two-dimensional mode at a low temperature of 300°C.[59] Shen et al. found a similar result for $In_xAl_{1-x}As$ step-graded buffers on GaAs.[58]

Use of low-temperature growth can also be helpful in reducing surface roughness associated with phase separation or clustering. Shang et al.[22] studied the influence of temperature and arsenic overpressure on the morphology in step-graded $In_xAl_{1-x}As$. They found that low-temperature growth at 380°C resulted in the smoothest morphology, with rms roughness of 0.802 nm as determined by AFM using a 5 × 5 μm observation area. In-As and Al-As have very different bond energies, which can lead to InAs and InAl clustering; Shang et al. attributed the improved smoothness at low temperature to the suppression of InAs-AlAs clustering. Increasing the As overpressure from 5.0×10^{-6} torr to 7.6×10^{-6} torr gave rise to a three-dimensional growth mode as seen by the spotty RHEED pattern, and worsened the surface roughness.

In some studies, it has been reported that growth of a thick, uniform layer on top of the step-graded buffer can reduce the roughness associated with cross-hatch, by burying the

misfit dislocations responsible for the cross-hatch. Balakrishnan et al.[37] used a 2 μm uniform layer of $Al_{0.5}Ga_{0.5}As_{0.54}Sb_{0.46}$ on top of a step-graded $Al_{0.5}Ga_{0.5}As_{1-y}Sb_y$ buffer for this purpose. Though the rms surface roughness was not given, it was amenable to the deposition of InAs quantum dashes.

6.6.4 Crystallographic Tilting in a Step-Graded Layer

Step-graded buffers typically exhibit crystallographic tilt, which accumulates at each step[31] and which has been reported to degrade the quality of device layers grown on top.[21] Lee et al.[31] studied crystallographic tilting in step-graded $In_xAl_{1-x}As$ buffer layers grown on GaAs substrates by MBE. The tilt increased with each step of the buffer, and its magnitude was consistent with the assumption that all misfit dislocations were 60° dislocations with the same sign tilt component (type II relaxation[60]). The tilt also decreased at the reverse step of the overshoot graded structure.

Gonzalez et al.[21] studied tilting in step-graded $In_xGa_{1-x}As$ buffers having two steps. Crystallographic tilt was present at each of the two interfaces, but the introduction of a graded transition zone at the second interface reduced it greatly. Based on this observation and the reduction of the misfit dislocation density with a transition zone, it was proposed that the difference comes about because there is a greater proportion of edge dislocations with the use of a graded transition, and edge dislocations with in-plane Burgers vectors do not introduce any tilt.

6.7 Linearly Graded Layer

Buffer layers with linearly graded composition, and therefore lattice constant, have been extensively investigated in several material systems, such as $GaAs_{1-x}P_x$/GaAs (001),[44,61,106,107] $In_xGa_{1-x}As$/GaAs,[4,62–71] $In_xAl_{1-x}As$/GaAs,[29,70,72–78] $In_xAl_yGa_{1-x-y}As$/GaAs,[34,36,59,79–82] $Si_{1-x}Ge_x$/Si,[2,83–85,96] $In_xGa_{1-x}P$/GaAs,[86–88] $In_xGa_{1-x}P$/GaP,[89] ZnS_ySe_{1-y}/GaAs,[4,90] and $In_xGa_{1-x}Sb$/GaSb.[91,99] A possible advantage of continuous grading is that layer-by-layer growth may be maintained without the intrusion of island growth associated with large, abrupt changes in composition.[88] Haupt et al.[36] compared the surface morphologies of $In_{0.52}Ga_{0.48}As$ HEMTs on linear and step-graded $In_xAl_{1-x}As$ buffer layers; as shown in the Nomarski micrographs of Figure 6.9, the structure with the linear buffer exhibited a cross-hatch pattern, whereas the step-graded device had rough, irregular morphology, which was attributed to three-dimensional growth. In another study involving quaternary $In_xGa_yAl_{1-x-y}As$ buffer layers, Haupt et al.[92] found that devices on linear buffers possessed smoother morphology and significantly higher 77K mobility (38,000 vs. 26,000 cm^2/Vs) than those on step-graded buffers.

6.7.1 Approaches to Linear Grading

Several approaches to linear compositional grading are shown in Figure 6.10, along with their approximate misfit dislocation density profiles. By far the most studied approach is the "forward-graded" buffer shown in Figure 6.10a, in which the lattice mismatch is zero at the substrate interface and increases linearly to match the lattice constant of the device

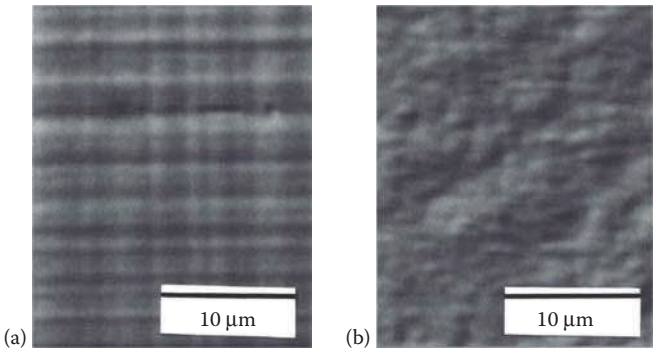


FIGURE 6.9
Nomarski micrographs of $In_{0.52}Ga_{0.48}As$ HEMTs on (a) linear and (b) step-graded $In_xAl_{1-x}As$ buffer layers. (Reprinted from M. Haupt et al., *Appl. Phys. Lett.*, 69, 412 [1996]. With permission. Copyright 1996, American Institute of Physics.)

layer,* and the lattice mismatch is given by $f = C_f z$. Although the case of tensile mismatch ($C_f > 0$) is shown in Figure 6.10a, the compressive case is simply a mirror image. In the overshoot graded layer shown in Figure 6.10b, the lattice mismatch in the graded buffer "overshoots" that in the device layer. This can compensate for incomplete relaxation in the graded layer, allowing in-plane lattice matching between the strained buffer layer and relaxed device layer.[5] The jump-graded buffer[96] of Figure 6.10c contains a nonzero interfacial mismatch, $f = f_0 + C_f z$, where f_0 and C_f both have the same sign as the device mismatch. It contains interfacial misfit dislocations, and can quickly reach a steady-state threading dislocation density, but dislocation interactions and pinning can be more pronounced than in the simple forward-graded layer. In the reverse-graded approach of Figure 6.10d, $f = f_0 + C_f z$, and f_0 has the same sign as the device mismatch, but C_f has the opposite sign. The relaxation behavior of reverse-graded layers is quite complex,[93] exhibiting force balance on grown-in dislocations at up to three critical layer thicknesses,[98] but preliminary work with this type of buffer indicates that the reverse grading may allow the bending over of threading dislocations associated with the interfacial mismatch, resulting in a low device threading dislocation density with a minimal buffer thickness.[94] In the following discussions, our focus will be the forward-graded approach illustrated in Figure 6.10a unless otherwise noted.

6.7.2 Critical Thickness in a Linearly Graded Layer

Fitzgerald et al.[95] have calculated the critical layer thickness for the onset of lattice relaxation in a linearly graded layer, using an approach similar to the Matthews energy derivation for an abrupt heterostructure. Suppose the distance from the interface is z and the lattice mismatch varies linearly with this distance so that $f = C_f z$, where C_f is the grading constant in cm^{-1}. At any distance from the interface, $f = \varepsilon_{\parallel} + \delta$, where $\varepsilon_{\parallel}$ is the in-plane strain and δ is the lattice relaxation. The dislocation dynamics and strain relaxation in a

* The "device layer" may actually include a series of layers, such as multiple quantum wells and graded layers, as in a graded index separate confinement heterostructure laser diode. For simplicity, we show the device layer as a single uniform layer with the same average lattice constant and total thickness as the actual device structure.

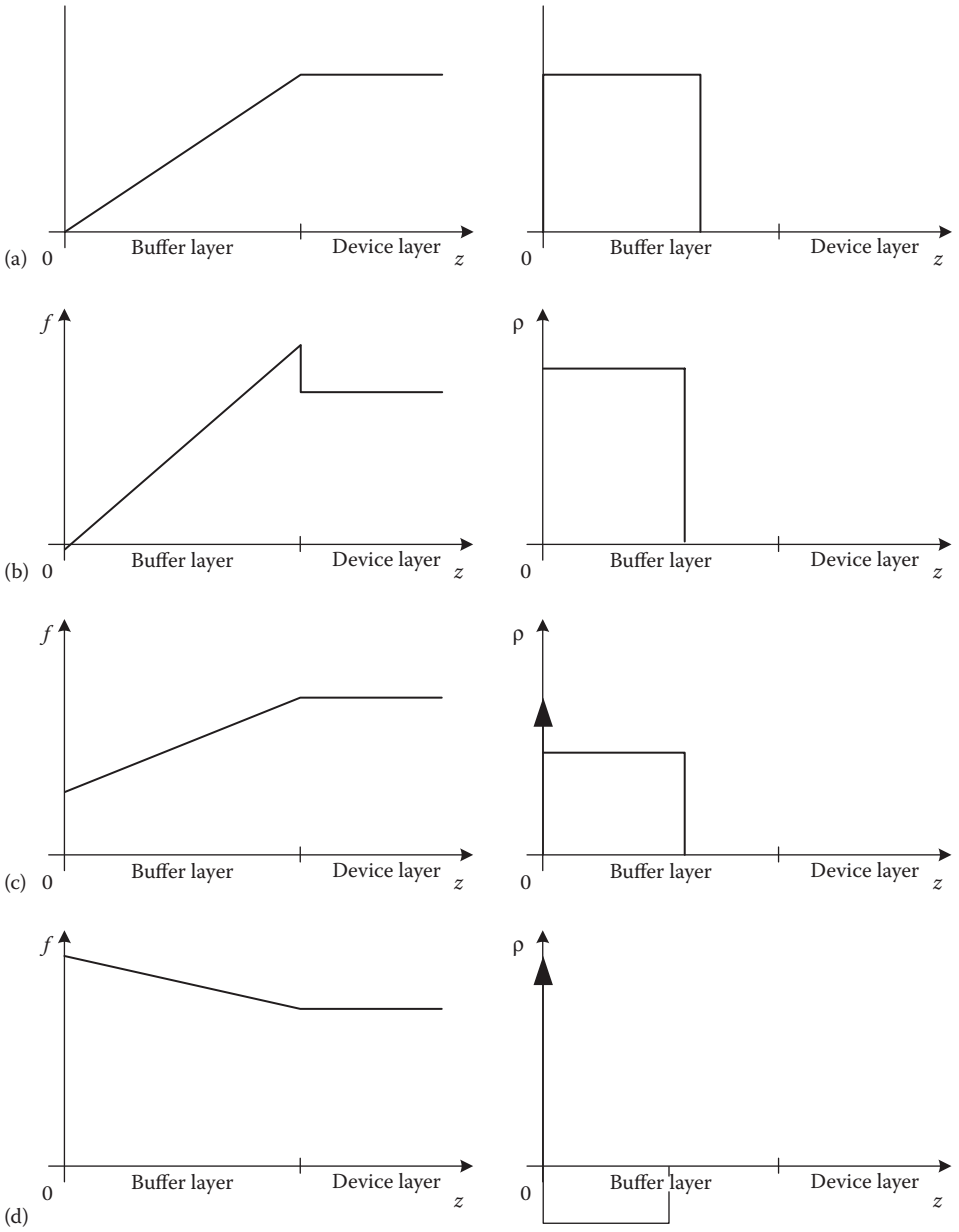


FIGURE 6.10
Approaches to linearly graded buffers: (a) forward graded, (b) overshoot graded, (c) jump graded, and (d) reverse graded. f is the lattice mismatch, ρ is the misfit dislocation density, and z is the distance from the substrate interface.

graded layer are rather complex, because the dislocations have distributed misfit components rather than well-defined misfit segments lying at or near the interface. However, the analysis can be greatly simplified by the assumption that $\varepsilon_{\parallel}$ and δ vary linearly with distance from the interface (as supported by experimental evidence), so that $\varepsilon_{\parallel} = C_{\varepsilon}z$ and $\delta = C_{\delta}z$. The elastic strain energy per unit area of total thickness h will therefore be

$$E_e = 2G\left(\frac{1+\nu}{1-\nu}\right)C_\varepsilon^2\int_0^h z^2 dz, \tag{6.36}$$

where G is the shear modulus and ν is the Poisson ratio. The line energy of misfit dislocations per unit area, assuming (001) heteroepitaxy of a diamond or zinc blende semiconductor with 60° misfit segments, will be

$$E_d = \frac{GbhC_\delta(1-\nu/4)}{\pi(1-\nu)}\left[\ln\left(\frac{h}{b}\right)+1\right], \tag{6.37}$$

where b is the length of the Burgers vector for the misfit dislocations. The critical layer thickness can be determined by $\partial(E_e + E_d)/\partial h = 0$, yielding

$$h_c^2 = \frac{3b(1-\nu/4)}{4\pi(1+\nu)C_f}\left[\ln\left(\frac{h_c}{b}\right)+1\right]. \tag{6.38}$$

Therefore, the critical thickness decreases monotonically with the grading coefficient.

6.7.3 Critical Layer Thicknesses in Linearly Graded Layers with Nonzero Interfacial Mismatch

For structures containing nonzero interfacial mismatch,[96] the behavior with respect to the critical layer thickness is much more complicated than that described in the previous section. In such layers, the lattice mismatch is given by $f = f_0 + C_f z$, where f_0 is the interfacial mismatch and C_f is the grading coefficient. If f_0 has the same sign as C_f, the layer can be referred to a "jump graded," whereas if f_0 and C_f have opposite signs, the layer can be referred to as "retro graded" or "reverse graded." A force balance model has been applied to find the critical layer thickness condition in graded layers of these types.

For the calculation of the critical layer thickness, it can be assumed that the initial relaxation occurs by the glide of threading dislocations inherited from the substrate to form interfacial misfit dislocations. The line tension in such a misfit segment has been calculated following the approach of Fitzgerald,[97] using the effective value of the Poisson ratio for (001) heteroepitaxy in the anisotropic case:

$$F_L = \frac{G_{\text{eff}}b\left(1-\nu\cos^2\alpha\right)}{2}\left[\ln\left(\frac{h}{b}\right)+1\right] \tag{6.39}$$

where:

- b is the length of the Burgers vector
- ν is the Poisson ratio (assumed to be constant throughout the structure)
- α is the angle between the Burgers vector and the line vector
- h is the layer thickness
- G_{eff} is the effective shear modulus for the interface, given by

$$G_{\text{eff}} = \frac{G_e G_s}{\pi(G_e+G_s)(1-\nu)}, \tag{6.40}$$

where:

G_e is the shear modulus for the epitaxial layer
G_s is the shear modulus for the substrate

For (001) heteroepitaxy of a zinc blende or diamond semiconductor, the Poisson ratio is given by

$$\nu(001) = \frac{C_{12}}{C_{11} + C_{12}}, \tag{6.41}$$

where C_{11} and C_{12} are the elastic stiffness constants (assumed constant throughout the structure). The glide force acting on the threading segment of the dislocation is

$$F_G = Yb\cos\lambda \int_0^h f(z)\,dz\,, \tag{6.42}$$

where:

λ is the angle between the Burgers vector and the line in the interface that is perpendicular to the intersection of the glide plane with the interface
Y is the biaxial modulus

Assuming (001) heteroepitaxy of diamond or zinc blende semiconductors,

$$Y = C_{11} + C_{12} - \frac{2C_{12}^2}{C_{11}}, \tag{6.43}$$

where C_{11} and C_{12} are the elastic stiffness constants for the epitaxial layer. Here, we have assumed that the elastic properties are approximately constant in the graded epitaxial layer. The condition for the onset of relaxation by the glide of grown-in threading dislocations is $F_G = F_L$. The Peierls force is not involved because this is an equilibrium condition. Therefore, the force balance condition is given by

$$\pm\frac{G_{\text{eff}}b\left(1-\nu\cos^2\alpha\right)}{2}\left[\ln\left(\frac{h}{b}\right)+1\right] = Yb\cos\lambda\int_0^h f(z)\,dz\,. \tag{6.44}$$

Integrating, the critical thickness or thicknesses are found from the solutions of

$$\frac{C_f h_c^2}{2} + f_0 h_c = \pm\frac{G_{\text{eff}}b\left(1-\nu\cos^2\alpha\right)}{2Y\cos\lambda}\left[\ln\left(\frac{h_c}{b}\right)+1\right]. \tag{6.45}$$

In order to make numerical calculations of the critical layer thicknesses for important cases of jump-graded and retro-graded layers, the approximate material properties can be assumed to correspond to Si (substrate) and $Si_{0.9}Ge_{0.1}$ (epitaxial layer). (These are summarized in Table 6.1.) However, the calculations given here will apply approximately to other material combinations having the same profiles of lattice mismatch strain.

Jump-graded layers exhibit a monotonic dependence of the critical layer thickness on the grading coefficient. This behavior is illustrated in Figure 6.11, which shows the critical

TABLE 6.1

Values Used to Calculate Critical Layer Thicknesses in Reverse-Graded and Overshoot Graded Layers

Parameter	Value
G_s	51 GPa
G_e	50 GPa
b	0.386 nm
ν	0.2635
C_{11}	156.4 GPa
C_{12}	56.2 GPa
α	60°
λ	60°

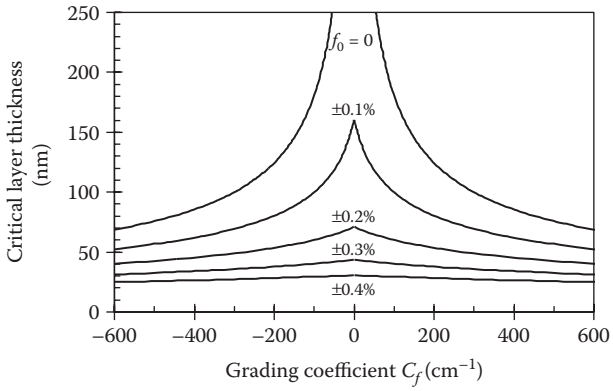


FIGURE 6.11
Critical layer thickness as a function of the grading coefficient, with pre-overshoot f_0 as a parameter, for jump-graded layers with sign(C_f) = sign(f_0). (Reprinted from J. E. Ayers, *Semicond. Sci. Technol.*, 23, 045018 [2008]. With permission. Copyright 2008, Institute of Physics.)

layer thickness for a jump-graded layer as a function of the grading coefficient and with the pre-overshoot f_0 as a parameter ($|f_0| = 0$, 0.1%, 0.2%, 0.3%, and 0.4%). For each value of the pre-overshoot f_0, the maximum value of the critical layer thickness is obtained with $C_f = 0$. As $|C_f|$ is increased, the absolute value of the average mismatch strain in the graded layer increases, so the critical layer thickness decreases. For a typical value of the pre-overshoot equal to ±0.1%, grading in the buffer layer may reduce its critical layer thickness by 50% or more. This is of significance in the design of graded structures, because it is usually necessary for the graded layer to be partially relaxed in order to affect the dislocation density in the top device layer. Depending on the growth rate and temperature, the graded buffer layer is typically grown to a thickness at least two times the critical layer thickness. Therefore, steeper grading may have the desirable effect of allowing a thinner buffer layer.

Retro-graded buffer layers exhibit much more complex behavior,[98] and the "critical layer thickness" may be multivalued for a particular combination of f_0 and C_f. This behavior is shown graphically in Figure 6.12, which shows the calculated critical layer thicknesses for retro-graded buffer layers as a function of the grading coefficient, with the interfacial mismatch equal to ±0.1% (Figure 6.12a), ±0.2% (Figure 6.12b), and ±0.4% (Figure 6.12c).

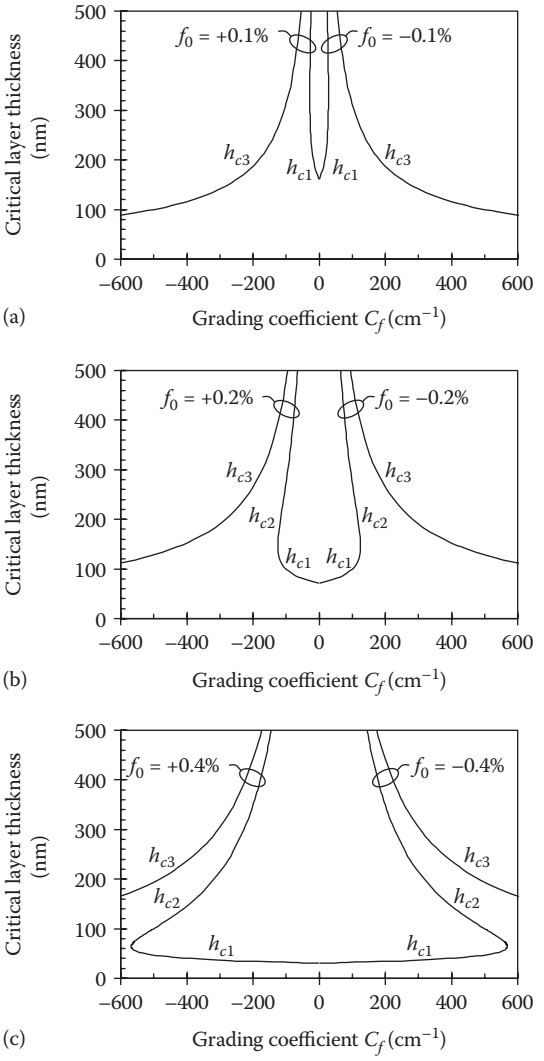


FIGURE 6.12
Critical layer thickness as a function of the grading coefficient, with pre-overshoot f_0 as a parameter, for retrograded layers with sign(C_f) = –sign(f_0). (a) f_0 = ±0.1%; (b) f_0 = ±0.2%; (c) f_0 = ±0.4%. (Reprinted from J. E. Ayers, *Semicond. Sci. Technol.*, 23, 045018 [2008]. With permission. Copyright 2008, Institute of Physics.)

Depending on the interfacial mismatch and grading coefficient, there may be up to three critical thicknesses.

This behavior may be understood by consideration of the forces acting on individual dislocations. For specificity, we can consider the forces on a grown-in dislocation on a (111) glide plane with a threading dislocation line vector of $[0\bar{1}1]$ and a Burgers vector $a/2[10\bar{1}]$, for the case of (001) heteroepitaxy of a diamond or zinc blende semiconductor. If this dislocation glides to create a length of misfit dislocation in the interface, it will glide in the $[1\bar{1}0]$ direction to relieve tensile mismatch strain, but in the $[\bar{1}10]$ direction to relieve compressive mismatch strain. The line tension force on the dislocation in the $[1\bar{1}0]$ direction is

$$F_L = \pm \frac{G_{\text{eff}} b\left(1 - \nu \cos^2 \alpha\right)}{2}\left[\ln\left(\frac{h}{b}\right) + 1\right] = F_L^+, F_L^- \tag{6.46}$$

where the sign depends on the direction of glide. The glide force on the grown-in dislocation in the $[1\bar{1}0]$ direction is given by

$$F_G = \Upsilon b \cos\lambda\left[f_0 h + \frac{C_f h^2}{2}\right]. \tag{6.47}$$

Consider first the case of $f_0 = +0.4\%$ and $C_f = -400\ \text{cm}^{-1}$ shown in Figure 6.13. $F_G = F_L^+$ at a thickness of $h_{c1} = 40$ nm (point A in Figure 6.13), and the glide of dislocations to relax mismatch strain is favorable for thicknesses greater than 40 nm. As the thickness of the retro-graded layer is increased, the glide force on the dislocation reaches a maximum (point B in Figure 6.13) at a thickness given by

$$h_{\text{crit}} = -\frac{f_0}{C_f}. \tag{6.48}$$

Beyond this thickness, the glide force actually reduces with layer thickness so that at a thickness of $h_{c2} = 146$ nm (point C in Figure 6.13) the line tension and glide forces balance once again. Lattice relaxation will therefore cease for $h > h_{c2}$. If the retro-graded layer is grown sufficiently thick, the glide force can change sign and the dislocation may glide in the negative direction ($[\bar{1}10]$ direction) if the thickness exceeds a value of $h_{c3} = 237$ nm (point D in Figure 6.13) for which $F_G = F_L^-$.

Qualitatively similar behavior can be observed for the case of $f_0 = +0.4\%$ and $C_f = -200\ \text{cm}^{-1}$, as illustrated in Figure 6.14. In this case, using the same assumptions as before, $h_{c1} = 17$ nm, $h_{c2} = 350$ nm, and $h_{c3} = 435$ nm. However, if the grading is made sufficiently steep, there is no thickness for which positive glide of the dislocation is favorable. This is shown in

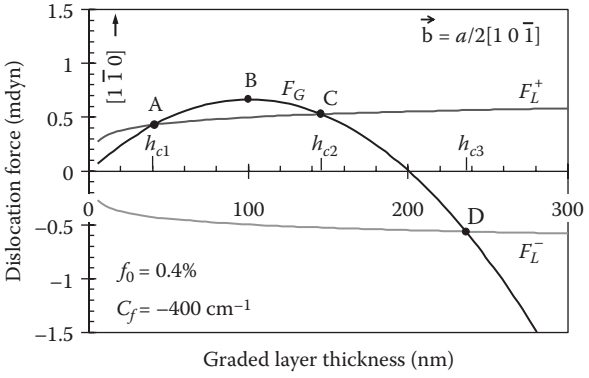


FIGURE 6.13
Forces acting on a grown-in dislocation in a retro-graded heteroepitaxial layer with $f_0 = +0.4\%$ and $C_f = -400\ \text{cm}^{-1}$. (001) heteroepitaxy of a diamond or zinc blende semiconductor was assumed. The grown-in dislocation is assumed to have a line vector of $[0\,\bar{1}\,1]$ in the substrate and glide on a (1 1 1) plane. The Burgers vector of the dislocation is assumed to be $a/2[1\,0\,\bar{1}]$. The assumed values of the material parameters were as given in Table 6.1. (Reprinted from J. E. Ayers, *Semicond. Sci. Technol.*, 23, 045018 [2008]. With permission. Copyright 2008, Institute of Physics.)

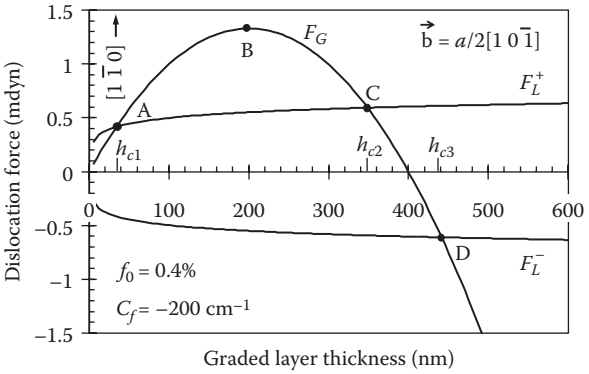


FIGURE 6.14
Forces acting on a grown-in dislocation in a retro-graded heteroepitaxial layer with $f_0 = +0.4\%$ and $C_f = -200\ \text{cm}^{-1}$. (001) heteroepitaxy of a diamond or zinc blende semiconductor was assumed. The grown-in dislocation is assumed to have a line vector of $[0\,\bar{1}\,1]$ in the substrate and glide on a (1 1 1) plane. The Burgers vector of the dislocation is assumed to be $a/2[1\,0\,\bar{1}]$. The assumed values of the material parameters were as given in Table 6.1. (Reprinted from J. E. Ayers, *Semicond. Sci. Technol.*, 23, 045018 [2008]. With permission. Copyright 2008, Institute of Physics.)

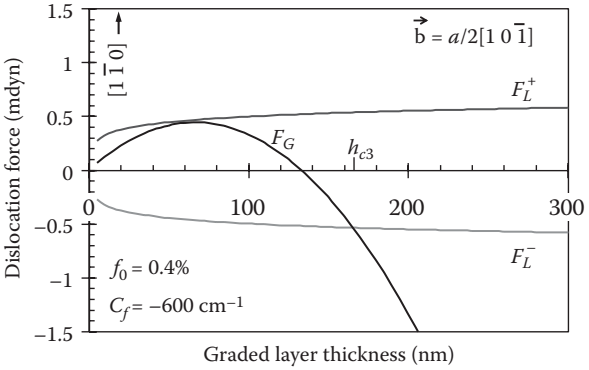


FIGURE 6.15
Forces acting on a grown-in dislocation in a retro-graded heteroepitaxial layer with $f_0 = +0.4\%$ and $C_f = -600\ \text{cm}^{-1}$. (001) heteroepitaxy of a diamond or zinc blende semiconductor was assumed. The grown-in dislocation is assumed to have a line vector of $[0\,\bar{1}\,1]$ in the substrate and glide on a (1 1 1) plane. The Burgers vector of the dislocation is assumed to be $a/2[1\,0\,\bar{1}]$. The assumed values of the material parameters were as given in Table 6.1. (Reprinted from J. E. Ayers, *Semicond. Sci. Technol.*, 23, 045018 [2008]. With permission. Copyright 2008, Institute of Physics.)

Figure 6.15 for the case of $f_0 = +0.4\%$ and $C_f = -600\ \text{cm}^{-1}$. Here, no solutions exist for h_{c1} or h_{c2}. The onset of lattice relaxation does not occur until the thickness exceeds $h_{c3} = 164$ nm, and is associated with negative glide of the dislocations (in the $[\bar{1}10]$ direction).

To summarize, modeling of the strain relaxation and dislocation dynamics in retro-graded buffer layers requires the consideration of the multiple critical layer thicknesses and the associated complications in the lattice relaxation process. It should also be emphasized that, in the preceding discussions, it has been assumed that the graded layer relaxes slowly enough so that the in-plane strain is approximately equal to the lattice mismatch

strain *f*. In cases for which significant lattice relaxation occurs, h_{c2} and h_{c3} become functions of the lattice relaxation.

6.7.4 Misfit Dislocations and Strain in a Linearly Graded Layer

Early experimental studies by Abrahams et al.[44] with vapor phase epitaxial $GaAs_{1-y}P_y$/GaAs (001) using XTEM revealed that the linearly graded material contained a nearly constant areal density of misfit dislocations ρ_A, that this density of misfit dislocations was proportional to the rate of grading, and that the surface density of threading dislocations *D* was also proportional to the rate of grading. Sacedon et al.[65] observed similar behavior in linearly graded MBE-grown $In_xGa_{1-x}As$/GaAs (001), but noted the existence of an MDFZ adjacent to the surface, above the dislocated region with nearly constant ρ_A.

Tersoff[5] developed an approximate model for the equilibrium misfit dislocation density and strain profiles in a linearly graded layer, based on a minimization of strain energy and dislocation line energy. It was assumed that the lattice mismatch, like the composition, varies linearly with distance from the interface and is given by

$$f(z) = C_f z, \tag{6.49}$$

where:

C_f is the grading coefficient
z is the distance from the interface

Tersoff showed that misfit dislocations are first introduced at a critical layer thickness given by

$$h_c = \sqrt{\frac{2F_d}{bYC_f \sin\alpha \sin\phi}}, \tag{6.50}$$

where F_d is the line energy of misfit dislocations and Y is the biaxial modulus,* which for a zinc blende semiconductor with (001) orientation is given by $Y = C_{11} + C_{12} - 2C_{12}^2/C_{11}$, where C_{11} and C_{12} are the elastic stiffness constants. Lattice relaxation is accompanied by the introduction of an approximately constant areal (cross-sectional) misfit dislocation density, which is just sufficient to relax the additional misfit introduced by grading:

$$\rho_A = \frac{C_f}{b \sin\alpha \sin\phi}. \tag{6.51}$$

The thickness of the dislocated region, in which this nearly constant misfit dislocation density exists, is given by

$$z_d = h - \sqrt{\frac{2F_d}{bYC_f \sin\alpha \sin\phi}}, \tag{6.52}$$

* Here the term *biaxial modulus* is used to refer to the Young's modulus for the case of biaxial stress. In other words, it is the ratio of the in-plane stress to in-plane strain for the condition of biaxial stress, $Y = \sigma_{\parallel}/\varepsilon_{\parallel}$, with $\sigma_{xx} = \sigma_{yy} = \sigma_{\parallel}$ and $\sigma_{xy} = \sigma_{yx} = \sigma_{yz} = \sigma_{zy} = \sigma_{zx} = \sigma_{xz} = 0$.

where h is the layer thickness. Above this, there is an MDFZ with thickness equal to h_c. In their XTEM study of linearly graded $In_xGa_{1-x}As/GaAs$ (001), Sacedon et al.[65] found the thickness of the MDFZ was close to that predicted by Tersoff.[5]

Bertoli et al.[4] used detailed numerical calculations to reveal small departures from the ideal rectangular profile for misfit dislocations found in the approximate analysis of Tersoff, and these are illustrated in Figure 6.16 for the case of linearly graded $Si_{1-x}Ge_x$/Si (001). First, there is a second MDFZ adjacent to the interface. This is because the lattice mismatch at the interface is zero, and introducing misfit dislocations there would actually

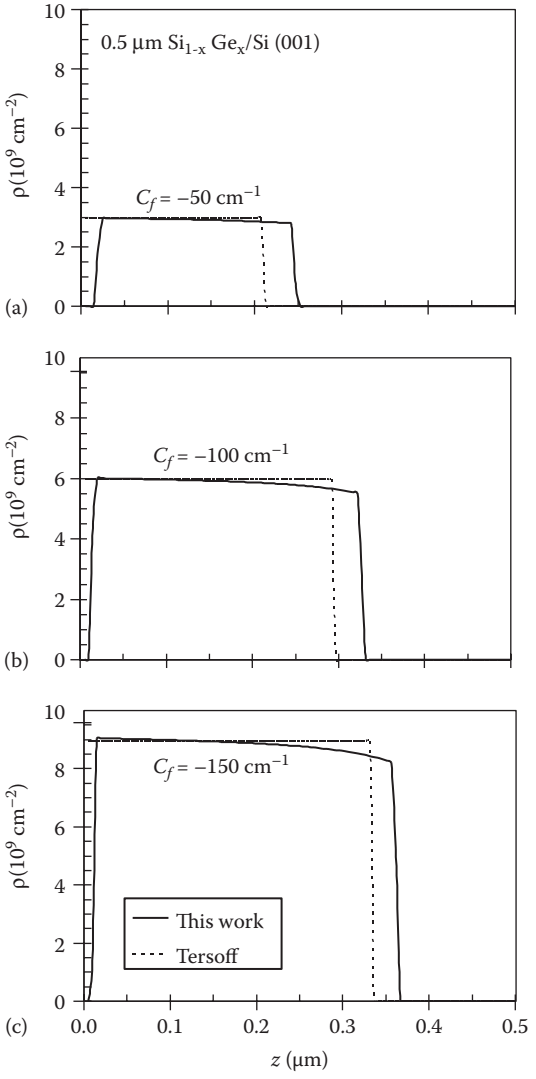


FIGURE 6.16
Calculated equilibrium misfit dislocation density profiles for 0.5 μm thick linearly graded layers of $Si_{1-x}Ge_x$-on-Si (001) substrates, with grading coefficients of (a) –50 cm^{-1}, (b) –100 cm^{-1}, and (c) –150 cm^{-1}. The dashed profiles were calculated using the approximate Tersoff model,[5] while the solid profiles were calculated using detailed numerical calculations by Bertoli et al.[4] (Reprinted from B. Bertoli et al., *J. Appl. Phys.*, 106, 073519 [2009]. With permission. Copyright 2009, American Institute of Physics.)

increase the strain energy. The thickness of the interfacial MDFZ is small in the linearly graded layer unless a very small grading coefficient is used, but has been observed by XTEM in $In_xGa_{1-x}As$ on GaAs.[17] Second, the misfit dislocation density is not constant, but tapered, in the dislocated region. The extent of this tapering may be up to 20% using typical material properties and grading coefficients. The underlying physical causes for this tapering are the change in the length of the Burgers vector with the variation in lattice constant, the inherent variation in the grading coefficient with a linear variation in lattice constant, and the reduction in the dislocation line energy with distance from the interface. In $In_xGa_{1-x}As$ and $Si_{1-x}Ge_x$ graded layers, the lattice constant increases with distance from the interface, so the misfit dislocation density decreases in the dislocated region. Additional differences are expected due to differences in the mobility of α and β dislocations, which cause asymmetries in the misfit dislocation densities along the [110] and $[1\bar{1}0]$ directions. This effect, not considered by Tersoff[5] or Bertoli et al.,[4] has been observed experimentally in linearly graded layers of $In_xGa_{1-x}As$/GaAs (001) grown by MBE.[69] In addition, Capotondi et al. showed that the thickness of the surface MDFZ is different for the misfit dislocations along the two ⟨110⟩ directions.[30]

The approximate equilibrium strain profile in the linearly graded layer was found by Tersoff[5] with the assumptions that (1) the strain is completely relaxed in the dislocated region and (2) the derivative of the strain in the MDFZ is equal to the grading coefficient, yielding

$$\varepsilon_{\|} = \begin{cases} 0, & z \le z_d; \\ C_f(z - z_d), & z > z_d. \end{cases} \tag{6.53}$$

Therefore, the linearly graded layer has a large built-in surface strain that helps to sweep out threading dislocations from the MDFZ, given approximately by

$$\varepsilon_{\|}(h) = \sqrt{\frac{2F_d C_f}{bY \sin\alpha \sin\phi}}. \tag{6.54}$$

Bertoli et al.[4] used detailed numerical calculations for $Si_{1-x}Ge_x$/Si (001) graded layers and found small departures from the approximate strain relationships given by Tersoff. The presence of the interfacial MDFZ introduces nonzero strain in the dislocated region, and also results in a somewhat larger built-in strain at the surface.

Implicit in Tersoff's derivation of the critical layer thickness for the linearly graded layer is the assumption that the dislocation line energy is independent of distance from the interface. Removing this assumption, and minimizing the total energy, Fitzgerald et al.[95] found the critical layer thickness as shown in the previous section. By applying a minimum energy approach, Sidoti et al.[2] determined the nonzero separation of the initial misfit dislocations from the interface to be

$$z_c = \frac{h_c}{2} - \sqrt{\frac{h_c^2}{4} - \frac{b(1-\nu)(1-\nu\cos^2\alpha)}{4\pi|C_f|(1+\nu)^2 \sin\alpha\sin\phi}}, \tag{6.55}$$

and this is the thickness of the interfacial MDFZ when the first misfit dislocations are introduced. It is important to note that the thickness of this MDFZ varies as the layer is grown; in general, the thickness of the interfacial MDFZ is z_1, and this is equal to z_c only at the critical layer thickness.

The state of strain and crystallographic tilting in a graded heterostructure may be analyzed using HRXRCs. Dynamical simulations are used to predict the rocking curve for the assumed structure, which is refined to obtain the best fit between the simulation and measured rocking curve.[46–48] In this way, the depth profiles of composition and strain may be determined indirectly. A limitation of this method has been the assumption of perfect crystals, which renders it applicable strictly to pseudomorphic structures. Recently, the dynamical theory has been extended to metamorphic structures, and can in principle allow the depth profiling of the dislocation density as well as the composition and strain.[49,50]

Reciprocal space maps may also be used to assess the state of strain and crystallographic tilting in a metamorphic buffer. As with HRXRC, the tilting is easily determined by using a symmetric reflection measured at two or more azimuths. Once the tilt is accounted for, the RSM shows the variations of the in-plane and out-of-plane lattice constants in the graded structure. For example, Figure 6.17 shows RSM for light-emitting diode (LED) structures on uniform 1 μm thick $InAs_ySb_{1-y}$ grown on GaSb using an intermediate linearly graded buffer of $InAs_xSb_{1-x}$ and measured around the 335 reflection of the GaSb (001) substrate.[99] The mole fraction of Sb in the top layer was 0.2 (Figure 6.17a and b) and 0.44 (Figure 6.17c and d). RSMs were measured at azimuths of 0° and 180° with respect to the [110] direction to allow cancellation of the crystallographic tilt. The abscissa and ordinate are $3/2a_{\parallel}$ and $5/a_{\perp}$, respectively, where $a_{\parallel}$ and $a_{\perp}$ are the in-plane and out-of-plane lattice constants, respectively. The red line shows the loci of 335 reflections from fully relaxed cubic material with varying lattice constants. The perpendicular blue lines represent constant in-plane lattice constants, coherently grown with the top of the linearly graded buffer. In each of the two structures, the linearly graded buffer gives rise to intensity along the diagonal line, indicating nearly complete relaxation. For the structure with 20% antimony, the uniform layer reciprocal lattice point, indicated by L, is off the diagonal line. From its position, the in-plane and out-of-plane lattice constants may be determined, and if the elastic constants are estimated, the relaxed lattice constant and state of strain may be determined.

In situ measurements, such as those made using a multibeam optical stress sensor (MOSS), make it possible to study the evolution of the average film stress. Multibeam laser illumination of the growing surface allows determination of the sample curvature and, indirectly, the average stress in the growing film. Figure 6.18 illustrates the application of MOSS to a linearly graded $In_xAl_{1-x}As/GaAs$ (001) structure grown by MBE with overshoot. Figure 6.18a shows the compositional profile, Figure 6.18b shows the stress–thickness product as a function of thickness, and Figure 6.18c illustrates the average stress as a function of layer thickness.[29] Initially, the average stress increases monotonically with thickness, corresponding to pseudomorphic growth. Once the layer starts to relax, at ~200 nm, the average stress begins to decrease. In the top of the linearly graded buffer, the average stress increases approximately linearly with thickness, indicating a near-constant strain in the top MDFZ of the graded buffer. The overshoot design,[59,62,100–102] frequently used in linearly graded metamorphic buffers, is intended to allow unstrained growth of the uniform layer, but the linear change in average stress during its growth shows that exact lattice matching was not achieved. However, MOSS monitoring could be used to help attain closer lattice matching in overshoot-graded structures.

Kujofsa and Ayers[103] compared the equilibrium configurations of step-graded and linearly graded metamorphic buffer layers (MBLs) in two material systems, $In_xGa_{1-x}As$ on GaAs (001) and $GaAs_{1-y}P_y$ on GaAs (001), using an energy minimization approach similar to that of Bertoli et al.[4] For the linearly graded case, the composition in the buffer layer was

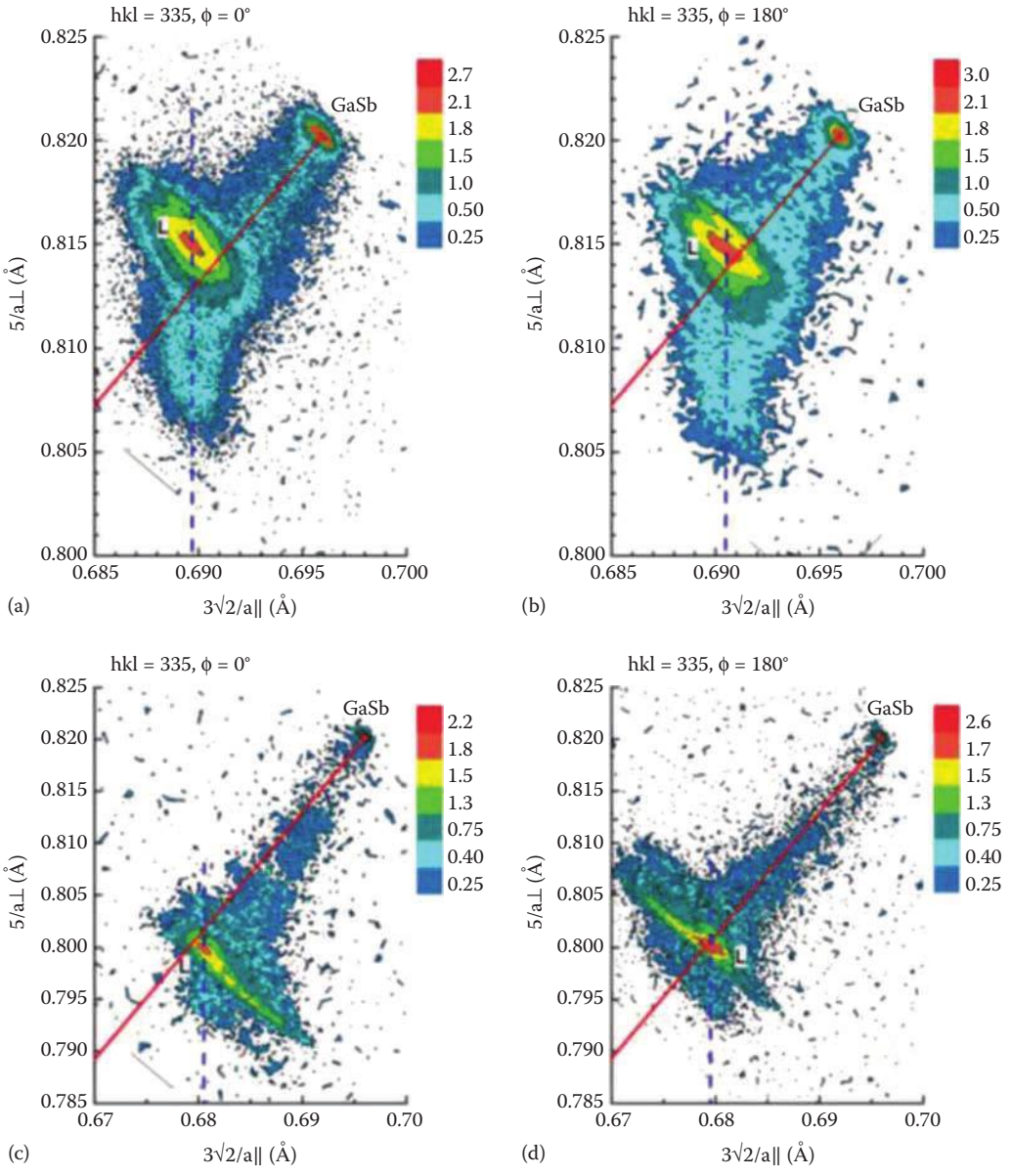


FIGURE 6.17
Reciprocal space maps for LED structures on uniform 1 μm thick $InAs_ySb_{1-y}$ grown on GaSb using an intermediate linearly graded buffer of $InAs_xSb_{1-x}$ and measured around the 335 reflection of the GaSb (001) substrate. The mole fraction of Sb in the top layer was 0.2 (a, b) and 0.44 (c, d). (Reprinted from G. Belenky et al., *Appl. Phys. Lett.*, 102, 111108 [2013]. With permission. Copyright 2013, American Institute of Physics.)

varied from lattice matched at the substrate interface to a composition of x_h (for $In_xGa_{1-x}As$) or y_h (for $GaAs_{1-y}P_y$) at the surface of the epilayer. The surface lattice mismatch varied from $f_h = 0.21\%$ to 2.2% in the structures considered. The step-graded layers utilized five steps of equal thickness, with equal steps in composition, so that the first layer was designed to have one-fifth of the lattice mismatch in the top layer.

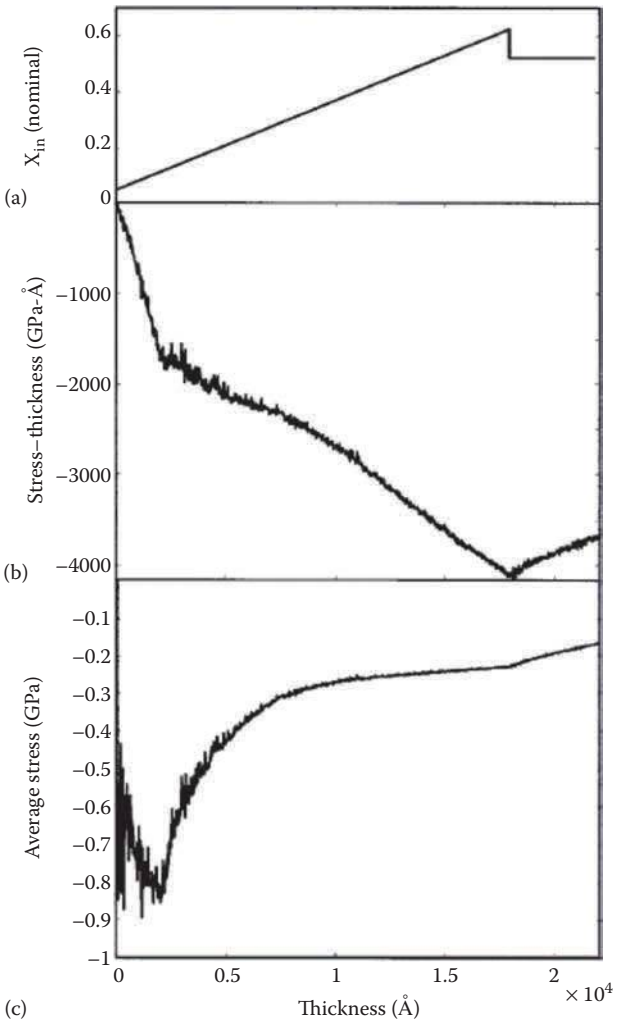


FIGURE 6.18
Application of the *in situ* MOSS characterization to a linearly graded InxAl1-xAs/GaAs (001) structure grown by MBE with overshoot: (a) compositional profile, (b) stress–thickness product as a function of thickness, and (c) average stress as a function of layer thickness. (Reprinted from C. Lynch et al., *J. Vac. Sci. Technol. B*, 22, 1539 [2004]. With permission. Copyright 2004, American Vacuum Society.)

In linearly graded structures, the misfit dislocation density could be approximated by a rectangular function, which was sandwiched between interfacial and surface MDFZs. The thickness of the surface MDFZ decreased as the total mismatch was increased, and varied from 320 nm at a mismatch of 0.64% to 410 nm at an ending mismatch of 1.48%. On the other hand, the thickness of the interfacial MDFZ was less than 10 nm for all cases investigated. The surface strain increases monotonically, but sublinearly, with the overall mismatch, as shown in Figure 6.27.

In the step-graded layers, misfit dislocations are introduced at discrete distances from the substrate interface corresponding to the steps in the layer. For a high value of ending mismatch, each step layer contains an MDFZ above a layer of interfacial dislocations and the MDFZ thickness is fixed by the step thickness. If the ending mismatch is decreased

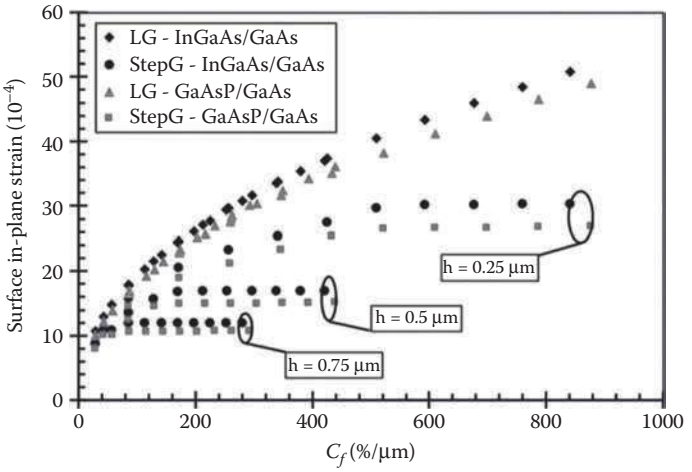


FIGURE 6.19
Surface in-plane strain as a function of the grading coefficient for step- and linearly graded InGaAs/GaAs (001) and GaAsP/GaAs (001). The thickness of the linearly graded layer varies from 0.25 μm to 0.75 μm. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst*., (24), 1520009 [2015]. With permission. Copyright 2016, World Scientific.)

sufficiently, interfacial dislocations are seen to be missing from one or more step layers adjacent to the free surface. For example, with an ending mismatch of 0.64% and a 100 nm step size, the width of surface MDFZ is 200 nm, or two steps. Typically, the surface MDFZ thickness exceeds that observed in a linearly graded layer with the same ending mismatch, resulting in a reduced surface strain compared with the linear case, as shown in Figure 6.19. GaAsP layers exhibit lower values of surface strain compared with InGaAs layers with equal grading coefficient and overall mismatch, and this difference was attributed to the greater elastic stiffness constants of GaAsP. As a first-order approximation, the strain energy in the two types of films is the same, with the lower strain compensating for the greater elastic constants in GaAsP.

Figure 6.20 shows the average misfit dislocation density as a function of the grading coefficient. The dependence could be fit approximately by exponential characteristics with different exponents corresponding to lower grading coefficients ($C_f < 200\%/\mu m$) and higher grading coefficients ($C_f > 200\%/\mu m$). The behavior is similar in GaAsP and InGaAs, showing that the difference in elastic constants is unimportant in this regard.

6.7.5 Threading Dislocations in a Linearly Graded Layer

The threading dislocation behavior of the linearly graded buffer is not completely understood at the present time. Nonetheless, experimental and modeling studies have shed some light on this important issue. In their pioneering study of linearly graded $GaAs_{1-y}P_y$/GaAs (001),[44,104–107] Abrahams et al.[44] made a number of important observations regarding the types, configurations, and densities of misfit and threading dislocations in linearly graded material. They found that the surface threading dislocation density D varied linearly with the grading coefficient C_f, over two orders of magnitude change, and devised a simple model to explain these findings. Three key concepts were developed in their paper: (1) that the misfit dislocations do not extend from one wafer edge to the other, but are segmented; (2) that there is a set of inclined threading dislocations associated with the segment ends

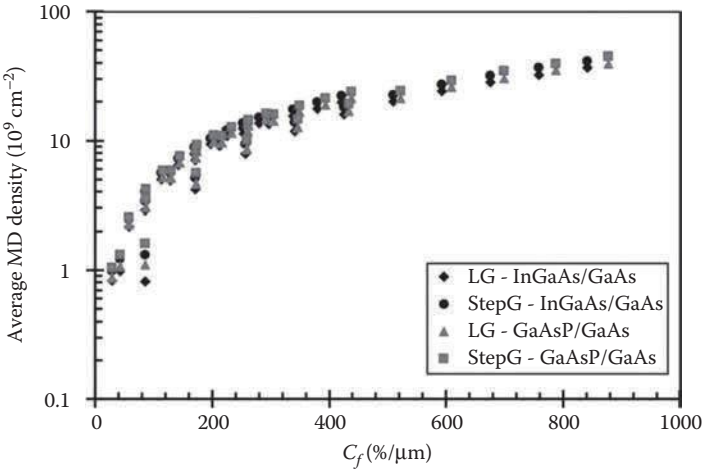


FIGURE 6.20
Average misfit dislocation density as a function of the grading coefficient for step- and linearly graded InGaAs/GaAs (001) and GaAsP/GaAs (001). (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 1520009 [2015]. With permission. Copyright 2016, World Scientific.)

because dislocations may not terminate within a crystal; and (3) that the growing graded layer reaches a steady-state condition in which new misfit segments are produced by the bending over of existing inclined dislocations rather than by the introduction of new ones. The cross-sectional density of misfit dislocations is

$$\rho_A = \frac{C_f}{b_\varepsilon}, \tag{6.56}$$

where b_ε is the misfit-relieving component of the Burgers vector. The number of misfit segments per unit volume will be

$$n_s = \frac{2\rho_A}{L_{\text{ave}}}, \tag{6.57}$$

where L_{ave} is the average length of misfit segments. If the density of inclined threading dislocations D reaches steady state after the growth of a thickness equal to the separation between misfit segments, $\rho_A^{-1/2}$, then the density will be

$$D = n_S \rho_A^{-1/2} = \frac{2C_f^{1/2}}{L_{\text{ave}} b_\varepsilon^{1/2}}, \tag{6.58}$$

where L_{ave} is the average length of misfit segments. If it is assumed that this average length is proportional to the average distance between misfit dislocations, $L_{\text{ave}} = m\rho_A^{-1/2}$, then

$$D = \frac{2C_f}{mb_\varepsilon}, \tag{6.59}$$

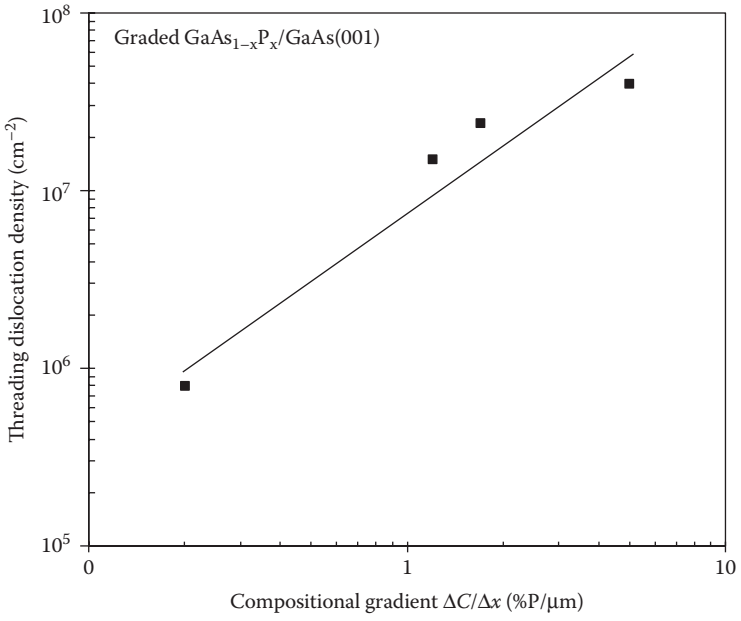


FIGURE 6.21
Threading dislocation density vs. compositional gradient for $GaAs_{1-x}P_x/GaAs$ (001) grown by vapor phase epitaxy. The grading coefficient is related to the compositional gradient by $C_f = \Delta f/\Delta z = 0.037\Delta C/\Delta z$, so that $\Delta C/\Delta x = 10\%/\mu m$ corresponds to $C_f = 0.37\%/\mu m$. (Adapted from M. S. Abrahams et al., *J. Mater. Sci.*, 4, 223 [1969]. With permission. Copyright 1969, Springer.)

and the density of threading dislocations will be proportional to the grading coefficient. This prediction was roughly verified by the experimental results of Abrahams et al., as shown in Figure 6.21. They found that the dislocation density increased in approximately linear fashion with the grading coefficient, from $D = 8 \times 10^5$ cm^{-2} with $C_f = 0.074\%/\mu m$, to $D = 4 \times 10^7$ cm^{-2} for $C_f = 0.185\%/\mu m$. By accounting for the fact that roughly half of the misfit dislocations were edge type while the other half were mixed 60° type, Abrahams et al. found a value of $m \approx 8$. Though this simple model is consistent with the experimental finding that the threading dislocation density is proportional to the grading coefficient in $GaAs_{1-y}P_y/GaAs$ (001), it does not include dislocation dynamics phenomena such as thermally activated glide, Peierls forces, dislocation interactions, and pinning. It therefore cannot explain the influence of temperature or the MDFZs. Moreover, it has been shown that the threading dislocation density increases strongly with the total lattice mismatch, even with the grading coefficient fixed, in $Si_{1-x}Ge_x/Si$ (001).[95]

Fitzgerald et al.[108] developed a dislocation dynamics model for a linearly graded metamorphic buffer grown at constant temperature. Key assumptions are that the layer is much greater than the critical layer thickness, that the threading dislocation density has reached a constant, steady-state value, and that dislocation glide is thermally activated with an exponential dependence on the effective stress.* Hence, the glide velocity for threading dislocations was assumed to be

* The effective stress is that component of the stress that is above and beyond the equilibrium value; $\tau_{eff} = \sigma_{\parallel} - \sigma_{eq}$, where $\sigma_{\parallel}$ is the actual in-plane stress and σ_{eq} is the equilibrium in-plane stress.

$$v = B\left(\frac{\tau_{\text{eff}}}{\sigma_0}\right)^m \exp\left(-\frac{U}{k_B T}\right), \tag{6.60}$$

where:

- B is a constant with units of velocity
- τ_{eff} is the effective stress
- σ_0 is a constant having units of stress
- m is a unitless constant generally between 1 and 2
- U is the activation energy for dislocation glide
- k_B is the Boltzmann constant
- T is the absolute temperature

If it is assumed that all dislocations are half loops and the misfit segments have 60° character, the rate of lattice relaxation is

$$\frac{d\delta}{dt} = \frac{\upsilon b D}{2} = \frac{bDB}{2}\varepsilon_{\text{eff}}^m \left(\frac{Y}{\sigma_0}\right)^m \exp\left(-\frac{U}{k_B T}\right), \tag{6.61}$$

where:

- b is the length of the Burgers vector
- Y is the biaxial modulus
- ε_{eff} is the effective in-plane strain (the actual in-plane strain minus the equilibrium value)

With the assumption of a constant rate of strain relaxation, the threading dislocation density may be solved for as

$$D = \frac{2gC_f \exp(U/k_B T)}{bB(Y/\sigma_0)^m \varepsilon_{\text{eff}}^m}, \tag{6.62}$$

where g is the growth rate. This model predicts a threading dislocation density proportional to the grading coefficient, similar to the Abrahams et al. model, but it also predicts a linear dependence on the growth rate.

Neither nucleation nor multiplication of dislocations was considered in the development of this model. However, it is possible that these processes can act to produce a steady-state threading dislocation density in the earlier stages of film growth. The model should then be applicable to the growth of the remaining thickness. Another key assumption of the dislocation dynamics model is that there are no impediments to the glide of the dislocations. Often, this is not the case. As will be demonstrated below, impediments to dislocation glide can drastically increase the defect densities in graded layers.

Fitzgerald et al.[108] applied the dislocation dynamics model to $In_xGa_{1-x}P$/GaP (001) graded layers by lumping the parameter B, the biaxial modulus, and the effective strain together in an adjustable constant C_1, yielding

$$D = \frac{2gC_f}{bC_1}\exp\left(\frac{U}{k_B T}\right). \tag{6.63}$$

This model was fit to experimental data for $In_xGa_{1-x}P$/GaP (001) graded layers grown to a final In composition of 10% by MOVPE in the temperature range 650°C–800°C. The growth

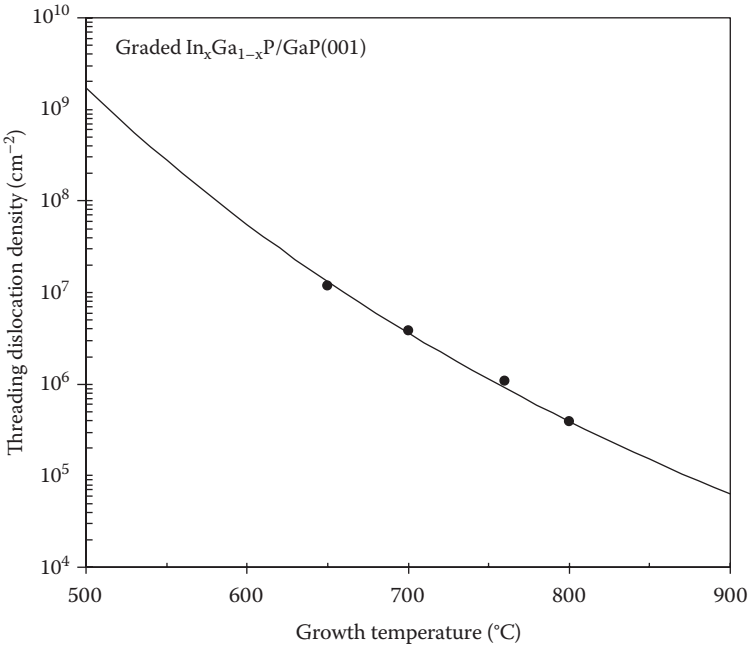


FIGURE 6.22
Threading dislocation density in graded $In_xGa_{1-x}P/GaP$ (001) as a function of growth temperature. All layers were graded to a final composition of 10% with a grading rate of 0.4%/μm (total thickness 1.9 μm). The filled circles represent experimental data. The curve was calculated using Equation 6.63 with $U = 2$eV, $g = 8.3 \times 10^{-4}$ μm s^{-1}, $C_f = 4 \times 10^{-3}$/μm, and $C_1 = 10^6$ cm s^{-1}.

rate was 3 μm h^{-1} (8.3×10^{-4} μm s^{-1}) and the grading coefficient was 0.4%/μm (4×10^{-3}/μm). Figure 6.22 shows the experimental data (filled circles) and the best fit based on Equation 6.63 using $U = 2$ eV and $C_1 = 10^6$ cm s^{-1}. The excellent fit between the model and the experimental data suggests the absence of impediments to dislocation glide in these graded buffer layers.

However, $In_xGa_{1-x}P/GaP$ (001) graded layers grown in the temperature range from 500°C–650°C exhibit a deterioration of the surface morphology and an anomalous increase in the threading dislocation density. This has been attributed to the occurrence of "branch defects," which can impede dislocation motion. Figure 6.23 shows PVTEM micrographs of $In_xGa_{1-x}P/GaP$ (001) graded layers with a top composition of 10% In grown by MOVPE at two different temperatures. The sample in Figure 6.23a, grown at 650°C, exhibits so-called branch defects, which are characterized by meandering lines of strain contrast. The threading dislocations appear to have segregated to the branch defects, indicating that the latter may be responsible for impeding the glide of the former. On the other hand, the sample of Figure 6.23b exhibits no visible branch defects or threading dislocation pileups. Given that the 650°C sample shows signs of dislocation pileups, it is surprising that its dislocation density lies so close to the curve in Figure 6.22. However, layers grown at still lower temperatures exhibit dislocation pileups to a greater degree and correspondingly higher threading dislocation densities.

At a fixed growth temperature of 760°C, the threading dislocation density in a graded $In_xGa_{1-x}P/GaP$ (001) buffer layer is a function of the final indium concentration, even with a constant grading coefficient. This is shown in Figure 6.24. Here, the filled squares

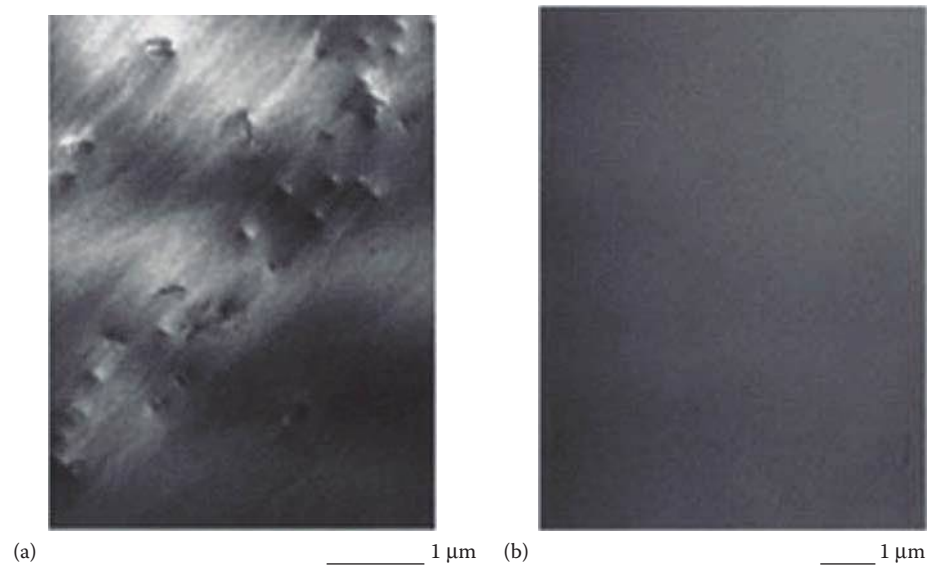


FIGURE 6.23
PVTEM micrographs of $In_xGa_{1-x}P/GaP$ (001) graded layers with a top composition of 10% In grown by MOVPE at two different temperatures: (a) 650°C and (b) 760°C. The branch defects in (a) appear to impede the glide of threading dislocations, but these are absent in (b). (Reprinted from E. A. Fitzgerald et al., *Mater. Sci. Eng. B*, 67, 53 [1999]. With permission. Copyright 1999, Elsevier.)

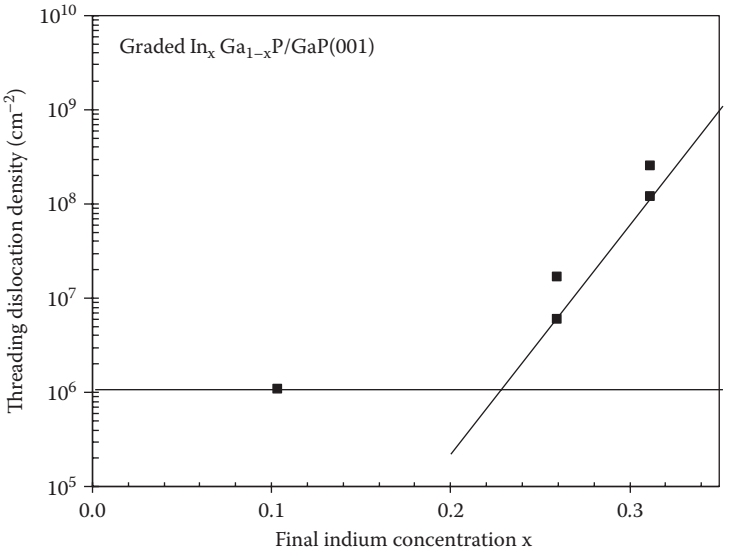


FIGURE 6.24
Threading dislocation densities in $In_xGa_{1-x}P/GaP$ (001) graded layers as a function of the final In composition, for layers grown with the same grading coefficient ($C_f = 4 \times 10^{-3}/\mu m$) and temperature (760°C).

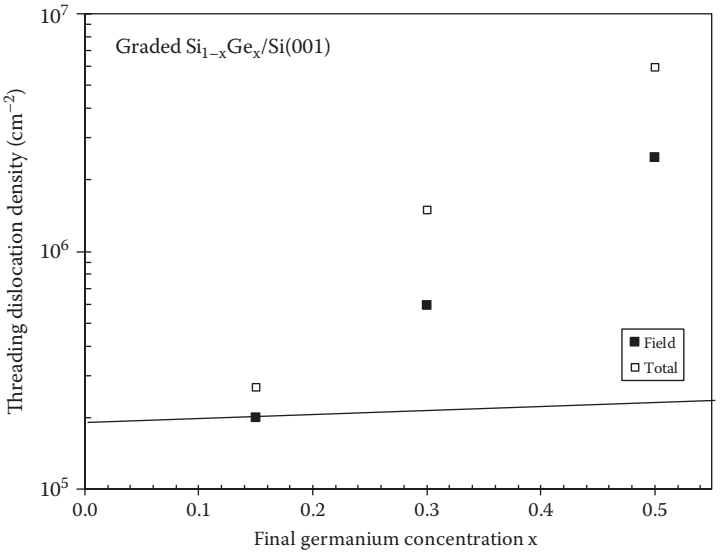


FIGURE 6.25
Threading dislocation densities in $Si_{1-x}Ge_x$/Si (001) graded layers as a function of the final Ge composition, for layers grown with the same grading coefficient ($C_f = 4.24 \times 10^{-3}$/μm) and temperature (700°C).[109] The total threading dislocation density includes the dislocations in the pileups. The field dislocation density is the threading dislocation density in the areas between the pileups.

represent measured threading dislocation densities for $In_xGa_{1-x}P$/GaP (001) graded layers, all of which were grown with the same grading coefficient ($C_f = 4 \times 10^{-3}$/μm) at temperature (760°C), but with different ending compositions. The sample with a final composition of x = 0.10 can be modeled using the fit of Figure 6.22, shown here with the flat line. Layers with a final indium composition of 0.2 or greater exhibit an anomalous increase in the dislocation density, which is thought to be related to lower average dislocation mobility (impediments to glide) associated with branch defects.

In the $Si_{1-x}Ge_x$/Si (001) system, it is also found that the dislocation density increases with the extent of the grading, as shown in Figure 6.25. Here, all of the samples were grown at the same temperature (750°C) and with the same grading coefficient ($C_f = 4.24 \times 10^{-3}$/μm). Here, the total threading dislocation density includes the dislocations in the pileups. The field dislocation density is the threading dislocation density in the areas between the pileups. The line was calculated using the dislocation dynamics model (Equation 6.62) with $g = 1.1 \times 10^{-3}$ μm s^{-1}, $C_f = 4.24 \times 10^{-3}$/μm, $U = 2.25$ eV, $B = 9.8 \times 10^3$ cm s^{-1}, $m = 2$, and $\varepsilon_{eff} = 1.33 \times 10^{-3}$. The biaxial modulus was estimated using the values for Si and Ge with a linear interpolation; this results in a slight upward slope of the line. The data point with a final composition of x = 0.15 can be fit with the dislocation dynamics model using this reasonable set of parameters. However, the layers graded to higher values of x (0.3 and 0.5) exhibit anomalous high threading dislocation densities. They also have a greater disparity between the total threading dislocation density and the field dislocation density. This indicates a greater tendency toward dislocation pileups, and an associated reduction in the effective strain, which can explain the elevated dislocation density. In the case of $Si_{1-x}Ge_x$, the pinning interactions between dislocations and the surface undulations have been cited as the sources of dislocation drag.

It has been argued on the basis of this model that linearly graded layers will have similar threading dislocation densities because it is not practical to vary the growth rate or the grading coefficient by orders of magnitude and still maintain a practical growth time.[108] An order of magnitude change in growth rate coupled with an order of magnitude variation in the grading coefficient could result in about two orders of magnitude change in the threading density, with all other parameters equal. However, it is becoming increasingly common to insert one or more low-temperature buffer layers in the structure, or to use temperature grading, step or nonlinear compositional grading, superlattice buffers, or some combination of these techniques. In such cases, the growth rate and grading rate can change significantly during growth, so that a much wider variation of surface dislocation density becomes achievable.

6.7.6 Crystallographic Tilting in a Linearly Graded Layer

Similar to step-graded buffers, linearly graded buffers tend to exhibit crystallographic tilt with respect to the substrate due to an imbalance of the populations of misfit dislocations with positive and negative tilt components. However, for the linear buffer the misfit dislocations are distributed throughout much of the thickness, so the tilting is expected to be continuous in nature rather than showing up as discrete events confined to step interfaces. In x-ray rocking curves, this shows up as broadening of the buffer layer diffraction profile, and in the reciprocal space map it exhibits as a diagonal spreading of the buffer diffraction contours.[31] Chyi et al.[20] compared the crystallographic tilting in linearly graded and step-graded $In_xGa_{1-x}As$ and $In_xAl_{1-x}As$. They found smaller tilts in the continuously graded layers, indicating a weaker imbalance between the slip systems. Lee et al. found a similar result in a comparison of step- and linearly graded $In_xAl_{1-x}As$ on GaAs (001).[31] Lee et al. also studied crystallographic tilting in MBE-grown single-slope and dual-slope linearly graded $In_xAl_{1-x}As$ buffers by RSMs; they found that the tilt increased superlinearly with indium composition, with sharp increases occurring at approximately $x = 0.6$.[31] Ihn et al.[59] used RSMs to show that a tilted low-temperature linearly graded $In_xAl_{1-x}As$ buffer could be brought back in line with the substrate axis by rapid thermal annealing (RTA) for 700°C for 30 s. Apparently the RTA reversed the imbalance in the dislocations with opposite tilt components; this implies that the imbalance in the populations is introduced by kinetic considerations, whereas minimum energy dictates nearly balanced dislocation populations.

In the heteroepitaxial system SiGe/Si (001), the observed tilts can be much greater in graded layers than in single heterostructures, for the same final composition. The measured tilts fall within the limits predicted for the cases of type I and type II relaxation[60] in both cases. However, the graded layers exhibit tilts much closer to the type II limit. This has been attributed to an anomalous strain relaxation mechanism,[109] which is unique to graded layers having high purity and low densities of surface defects.

With type II relaxation, the misfit dislocations along each [110] direction are introduced only by the most stressed slip systems (MSSSs). Therefore, in the graded layers exhibiting large tilts, there is a lattice relaxation mechanism that essentially excludes all but these MSSSs.

LeGoues, Mooney, and Chu[109] developed a model for the epitaxial layer tilt in linearly graded layers exhibiting this anomalous strain relaxation mechanism, which they called a modified Frank–Read (MFR) mechanism.[110] The underlying assumptions relating the tilt to the lattice relaxation by the eight slip systems are the same as in the Ayers, Schowalter, and Ghandhi model; however, the individual values of δ_i are assumed to be limited by dislocation nucleation rather than glide.

TABLE 6.2

Relationship between the Slip Systems Used by LeGoues, Mooney, and Chu[110] and Those Defined by Ayers, Ghandhi, and Schowalter[60]

MFR System	Slip Systems
(LeGoues, Mooney, and Chu)	**(Ayers, Ghandhi, and Schowalter)**
MFR1	S_3, S_5
MFR2	S_1, S_7
MFR3	S_4, S_8
MFR4	S_2, S_6

In their model, LeGoues, Mooney, and Chu grouped the eight active 60° slip systems for (001) heteroepitaxy in pairs, each of which is an MFR system. By this lattice relaxation mechanism, "corner dislocations" are associated with the simultaneous glide of two orthogonal dislocation segments on different {111} planes. Hence, the MFR1 system involves two dislocation segments, one from the 60° glide system S_3 and another from the glide system S_5. Similarly, MFR2 involves corner dislocations made of segments from S_1 and S_7. The MFR systems as defined by LeGoues, Mooney, and Chu are related to the slip systems tabulated by Ayers, Ghandhi, and Schowalter in Table 6.2.

By the MFR mechanism, the nucleation of a new dislocation produces one segment along each of the two orthogonal <110> directions. If the orthogonal segments always remain equal in length, then the MFR mechanism will result in equivalent strain relaxation in the two <110> directions.[109] In developing their model, LeGoues, Mooney, and Chu assumed this to be true, and that the miscut of the substrate introduced a change in the activation energy Δ for the nucleation of dislocations on the MSSS.

As a specific example, consider a (001) substrate inclined toward the [100]. (The axis of rotation associated with the substrate miscut is [010].) This should result in an epitaxial layer tilt about the [010] axis, requiring an imbalance between the MFR1 and MFR2 systems, but not between the MFR3 and MFR4 systems. Therefore, the numbers of dislocations in the four MFR systems are assumed to be such that

$$N_3 = N_4,$$

$$N_1 = N_3 \exp\left(-\Delta/k_B T\right),$$

and

$$N_2 = N_3 \exp\left(\Delta/k_B T\right), \tag{6.64}$$

where Δ is the change in nucleation energy arising from the miscut. The total number of dislocations is the sum

$$N_T = N_1 + N_2 + N_3 + N_4. \tag{6.65}$$

The imbalance in lattice relaxation by MFR1 and MFR2 results in the tilt so that

$$\Delta\Phi = \tan^{-1}\left[b_{\text{tilt}} N_{\text{tilt}}\right] = \tan^{-1}\left[b_{\text{tilt}}\left(N_1 - N_2\right)\right], \tag{6.66}$$

where b_{tilt} is the tilt component of the Burgers vector for MFR1 and MFR2. Finally, the ratio of the total dislocation density to the number producing tilt is expected to be

$$\frac{N_T}{N_{\text{tilt}}} = \frac{1+\cosh\left(-\Delta/k_BT\right)}{\sinh\left(-\Delta/k_BT\right)}. \tag{6.67}$$

Here, due to the exponential dependence, any appreciable change in the nucleation energy (e.g., $\Delta \approx -3k_BT$) will tend to drive the above ratio to 1, which is the type II limit.

This type of behavior is expected for any graded layer exhibiting the MFR mechanism of lattice relaxation. The MFR mechanism is believed to be active in graded SiGe/Si (001) and also graded InGaAs/GaAs (001), when the layers are of high purity and the substrate surfaces are relatively free from defects. In both material systems, dislocation loops have been observed to propagate deep into the substrate, and these substrate dislocations have been identified as a signature of the MFR mechanism. It is possible that the MFR mechanism is active in other heteroepitaxial material systems involving diamond or zinc blende semiconductors. However, this mechanism can only operate with low dislocation densities, and does not appear to be active in abrupt heterostructures.

The dependence of the nucleation energy on the substrate inclination is poorly understood at the present time, and this hinders the theoretical estimation of Δ. Therefore, it is not possible to know if the lattice relaxation, and therefore the crystallographic tilting, will be dominated by glide or nucleation of dislocations *a priori*. On the other hand, if it is assumed that the tilt is governed by nucleation, then the measured tilt $\Delta\Phi$ and dislocation density N_T can be used to estimate the change in activation energy Δ using the above equations. Such calculations have been made for graded SiGe grown on Si (001).[111]

In summary, it is now well established that tilting of heteroepitaxial layers is affected by substrate surface steps in strained heteroepitaxial layers.[112] In relaxed (or partly relaxed) heteroepitaxial layers, both the steps at the interface and the misfit dislocations[60] may contribute to the crystallographic tilting of the heteroepitaxial layer. It is generally accepted that net tilt results from an imbalance in the dislocation populations on the various slip systems.[60] The underlying cause for this imbalance is not entirely clear, but may relate to imbalances in the glide, multiplication, or nucleation of misfit dislocations. It is possible, in fact, that all three phenomena contribute to the dislocation imbalance (and hence the tilt) under certain conditions, depending on the material system and the growth conditions. It is likely that glide and multiplication of dislocations dominate the relaxation process and the tilt in most heteroepitaxial systems. However, nucleation may be the governing phenomenon in some compositionally graded systems that relax by an MFR mechanism, such as graded layers of SiGe/Si (0001). Further work, both theoretical and experimental, is needed to clarify this behavior.

Most of the work to date has been directed at diamond and zinc blende semiconductors. However, it has also been shown that the crystallographic tilts in relaxed AlN/6H-SiC (0001) can be predicted by the Nagai model for pseudomorphic layer.[112] This shows that the misfit dislocations in this material system do not contribute to the tilt. GaN on sapphire (0001) behaves similarly for small values of the substrate inclination. Preliminary results show that, with larger offcut angles, the presence of larger steps alters the tilting in this heteroepitaxial system. More experimental results are needed to characterize the tilting behavior of wurtzite semiconductors under a variety of conditions with various substrates. This will provide a better understanding of the mechanisms involved in the tilting of the materials, and therefore their relaxation mechanisms.

6.7.7 Surface Roughening and Cross-Hatch in a Linearly Graded Layer

Linearly graded buffers often exhibit strain-induced surface roughening that may take the form of corrugations,[100] cross-hatch,[63] or sometimes irregular surfaces. Cross-hatch is associated with misfit dislocations, and may be present in high-quality material with threading dislocation densities below the detection limit by PVTEM.[70] The roughness may be quite different along the [110] and [$\bar{1}$10] directions; for HEMT structures on GaAs substrates with linearly graded $In_xAl_{1-x}As$ buffer layers, Cordier et al.[102] measured rms surface roughness of 1.2–1.6 nm along the [110] direction, but 3.2–4.5 nm along the [$\bar{1}$10] direction.

The surface roughness associated with cross-hatch is strongly affected by the temperature, grading rate, and even doping in the metamorphic buffer. These effects are governed by strain-field-induced surface roughening, as described previously, but also by indium segregation in the case of $In_xGa_{1-x}As$ or $In_xAl_{1-x}As$ buffers. The increased indium concentration on the growth front has been modeled by Muraki et al.[113] as

$$x_{\text{seg}} = \frac{ahC_f R}{1-R}, \tag{6.68}$$

where:

- a is the lattice constant
- h is the buffer thickness
- C_f is the grading coefficient
- R is the segregation constant

This model predicts increased indium segregation, and therefore surface roughness, with increasing thickness or grading coefficient. Low-temperature growth decreases the surface roughness,[36,63,68,74,100,114] and this is expected because the reduced surface mobility of adatoms would decrease surface roughening by either mechanism. Haupt et al.[36] measured the rms roughness and 2DEG mobility for $In_{0.53}Ga_{0.47}As/In_{0.53}Al_{0.47}As$ HEMT structures grown on GaAs with linear and step-graded $In_xAl_{1-x}As$ buffers, for different buffer growth temperatures. Reduced buffer growth temperature, in the range 300°C–400°C, resulted in reduced roughness (Figure 6.26) and improved 77 K mobility (Figure 6.27). Song et al. investigated the influence of doping on surface roughness in low-temperature linearly graded $In_xGa_{1-x}As$ buffers on GaAs.[63] Beryllium doping reduced the surface roughness relative to undoped material, whereas silicon doping increased the surface roughness, as shown in Figure 6.28. These differences were proposed to be due to the influence of dopants on the segregation of indium; whereas silicon enhances this segregation, beryllium has the opposite effect. Surfactants can also modify the adatom mobility and reduce the surface roughness, as in the self-surfactant effect observed with antimony-containing graded buffers.[115]

6.7.8 Dual-Slope and Tandem Graded Layers

Several unique graded buffer layers have been created by the use of dual-slope buffers or buffers containing two different alloy semiconductors (tandem buffers).

Capotondi et al. used a dual-slope step-graded $In_xAl_{1-x}As$ buffer for the realization of $In_xGa_{1-x}As/In_yAl_{1-y}As$ quantum wells on GaAs.[30] In this buffer, the average grading coefficient was 510 cm^{-1} (5.1%/μm) in the first 600 nm and 310 cm^{-1} in the remaining 600 nm, with 50 nm thick steps. Because of the small step thickness, this buffer was considered

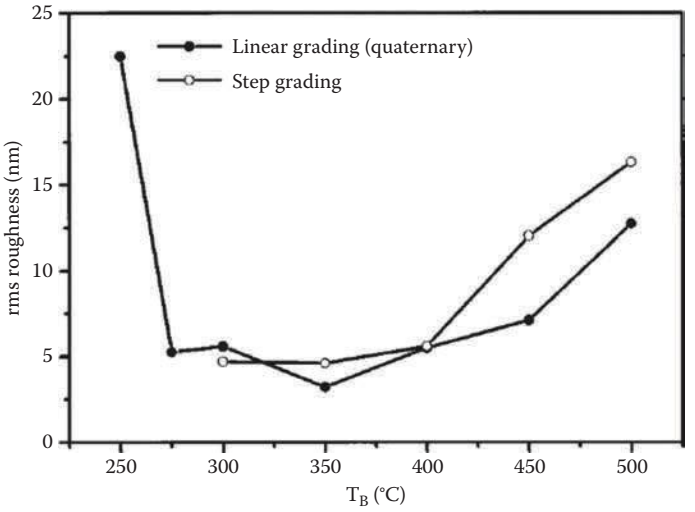


FIGURE 6.26
Root mean square surface roughness as a function of the buffer growth temperature, for $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures grown on GaAs substrates with linearly graded or step-graded $In_xAl_{1-x}As$ buffer layers. (Reprinted from M. Haupt et al., *Appl. Phys. Lett.*, 69, 412 [1996]. With permission. Copyright 1996, American Institute of Physics.)

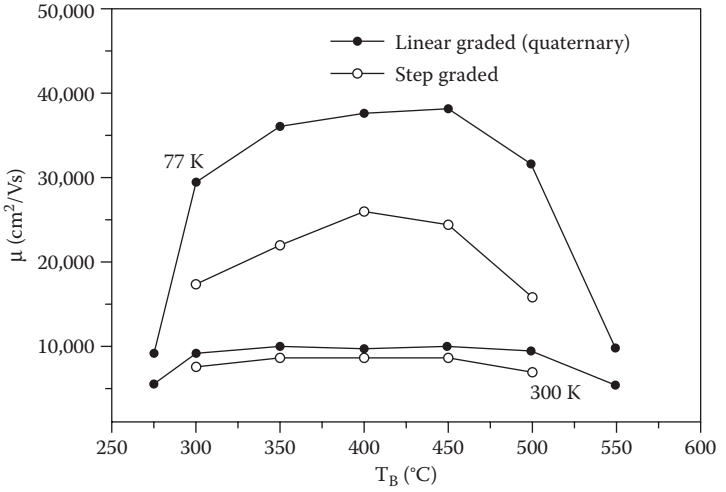


FIGURE 6.27
2DEG mobility as a function of the buffer growth temperature, for $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures grown on GaAs substrates with linearly graded or step-graded $In_xAl_{1-x}As$ buffer layers. (Reprinted from M. Haupt et al., *Appl. Phys. Lett.*, 69, 412 [1996]. With permission. Copyright 1996, American Institute of Physics.)

to approximate a continuous linearly graded layer, and the Tersoff model[5] was applied to analyze its behavior. XTEM micrographs showed that the material with the smaller grading coefficient had a lower misfit dislocation density, and that there was a 450 nm thick MDFZ adjacent to the surface.

Lee et al. used RSMs to investigate dual-slope continuously graded $In_xAl_{1-x}As/GaAs$ buffer layers, in which the top portion of each buffer had a larger grading coefficient

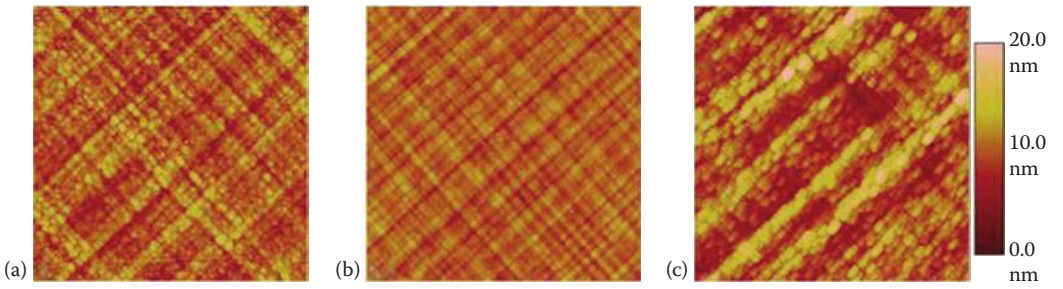


FIGURE 6.28
Cross-hatch morphology in linearly graded $In_xGa_{1-x}As/GaAs(001)$ grown by MBE. (a) Undoped, (b) Be doped, (c) Si doped. (Reprinted from Y. Song et al., *J. Appl. Phys.*, 106, 123531 [2009]. With permission. Copyright 2009, American Institute of Physics.)

than the bottom portion.[31] The initial slope in all layers was ~300 cm^{-1}, but the secondary slope was varied. They found that the buffers with a greater secondary slope exhibited reduced strain and crystallographic tilting but increased full width at half maximum (mosaic spread). Single-slope and dual-slope buffers showed similar behavior, which was controlled by the final grading coefficient. Therefore, in a dual-slope graded buffer, the mosaic spread may be controlled by adjusting the grading coefficient in the topmost portion of the buffer.

Yang et al. developed a tandem graded buffer using MOVPE-grown $In_xGa_{1-x}As$ plus $In_xGa_{1-x}P$ for InP-based devices on GaAs substrates.[71] The tandem design was used to avoid the problem of phase separation, which leads to surface roughening and high threading dislocation densities in $In_xGa_{1-x}As$ with high indium. The indium mole fraction was graded to 0.3 in the $In_xGa_{1-x}As$, and to 1.0 in the $In_xGa_{1-x}P$. The $In_xGa_{1-x}As$ buffer was grown at reduced temperature (450°C instead of 750°C) for compositions in the range $0.1 \leq x \leq 0.3$ to further suppress phase separation. By this approach, they obtained metamorphic InP on GaAs with a reported threading dislocation density of 7.9×10^6 cm^{-2} and an rms surface roughness of 7.4 nm on an AFM image of 40×40 μm.

6.8 Sublinearly and Superlinearly Graded Layers

Nonlinear graded buffer layers have been investigated in order to manipulate the MDFZs and their built-in strains, and also to control the steady-state threading dislocation density. The thickness of the MDFZ adjacent to the device may be enhanced by tapering the grading coefficient near the device (sublinear grading), while the thickness of the interfacial MDFZ may be enhanced by tapering the grading coefficient near the substrate (superlinear grading).

The role of the MDFZ adjacent to the device layer in promoting the glide of threading dislocations for the achievement of the longest possible misfit segments is well established.[5] It is desirable to maintain a thick MDFZ with large residual strain for maximum effectiveness. Modeling and experimental studies have shown that the thickness of this MDFZ may diminish upon growth of the device layer or structure on top of the graded buffer;[4] in other words, there is an undesirable "loading effect" of the device structure on

the buffer. In some cases, this MDFZ may disappear entirely due to loading. One solution to the problem is to overshoot the average lattice constant of the device layer in the graded buffer, so the in-plane lattice constant of the relaxed device structure matches that of the strained buffer layer. If the device structure is relaxed, it applies no stress to the buffer and the loading effect vanishes. However, this is not applicable in cases where it may be desirable to have built-in strain in the device structure to enhance carrier mobility or shift the emission or absorption wavelength. Another approach is to alter the grading profile to make it less susceptible to loading, and this can be achieved using sublinear grading adjacent to the device layer.[116]

Sublinear grading may also taper the steady-state threading dislocation density to a lower value. The models of Abrahams et al.[44] and Fitzgerald et al.[108] show that the steady-state threading dislocation density is proportional to the grading coefficient, so a gradual reduction of the grading coefficient in the sublinear buffer can help remove threading dislocations from the top of the buffer. This idea is reinforced by the observation that the threading dislocation density can be controlled by the grading coefficient near the top of a tandem graded buffer.[31]

There have been experimental and modeling studies of sublinear continuous grading in the material systems $In_xGa_{1-x}As/GaAs$[77,117,118] and $Si_{1-x}Ge/Si$,[4,116] and sublinear step grading has been investigated in the material system $InAs_yP_{1-y}/InP$.[45]

6.8.1 Critical Layer Thickness in Sublinear Exponentially Graded Layers

The use of sublinear grading alters the critical thickness compared with the linearly graded case. Sidoti et al.[117] considered force balance on a grown-in dislocation to determine the critical layer thickness in a buffer with a sublinear exponential lattice mismatch profile given by

$$f = f_\infty(1 - e^{-z/\gamma}), \tag{6.69}$$

where:

- f_∞ is the limiting mismatch
- z is the distance from the interface
- γ is the grading length constant

The approximate critical layer thickness for such a layer may be determined using a simple force balance model, as illustrated in Figure 6.1. Following the approach presented by Matthews and Blakeslee[1] for a structure containing uniform strained layers, it was assumed that a grown-in dislocation will elongate at the interface to create a length of misfit dislocation once the glide force F_G exceeds the dislocation line tension F_L. The glide force acting on a grown-in dislocation in the coherently strained layer on a (001) diamond or zinc blende substrate with $\varepsilon_{\|}(z) = f(z)$ is

$$F_G = \frac{2Gb(1+\nu)\cos\lambda}{(1-\nu)} \int_0^h f(z)dz, \tag{6.70}$$

where:

- G is the shear modulus, $G = (C_{11} - C_{12})/2$
- b is the length of the Burgers vector

- ν is the Poisson ratio, $\nu = C_{12}/(C_{11} + C_{12})$, for the (001) orientation
- λ is the angle between the Burgers vector and the line in the interface plane that is perpendicular to the intersection of the glide plane with the interface
- h is the layer thickness

For the case of a sublinear (convex-up) exponentially graded layer, the glide force is

$$F_G = \frac{2Gb(1+\nu)\cos\lambda}{(1-\nu)}\int_0^h f_\infty\left[1-e^{-z/\gamma}\right]dz = \frac{2Gbf_\infty(1+\nu)\cos\lambda}{(1-\nu)}\left[h+\gamma\left(e^{-z/\gamma}-1\right)\right]. \quad (6.71)$$

The line tension in the misfit segment that opposes the glide of the grown-in dislocation is

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h}{b}\right)+1\right], \quad (6.72)$$

where α is the angle between the Burgers vector and the line vector for the dislocations. Equating the glide and line tension forces at the critical layer thickness h_c yields

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)\left[\ln\left(h_c/b\right)+1\right]}{8\pi\left|f_\infty\right|(1+\nu)\cos\lambda}+\gamma\left(1-e^{-h_c/\gamma}\right), \quad (6.73)$$

which may be solved numerically. It is important to point out that this relationship is based on the assumption that misfit dislocations are first introduced at the substrate interface. Although a good assumption for a uniform strained layer, it will be subject to some degree of error if the first misfit dislocations are introduced away from the interface. Moreover, effective values for b and ν have been assumed, even though these parameters vary with composition in a graded layer.

The critical layer thickness may be determined more accurately by consideration of energy minimization, without invoking effective or average values of b and ν. If the graded layer is sufficiently thick to be partly relaxed, and therefore contains misfit dislocations of cross-sectional density $\rho(z)$, the in-plane strain is relaxed to

$$\varepsilon_{\|}(z) = f(z) + b'\int_0^z \rho(\xi)\,d\xi, \quad (6.74)$$

where b' is the misfit-relieving component of the Burgers vector parallel to the interface. The strain energy per unit area is

$$E_\varepsilon = \int_0^h Y\varepsilon_{\|}^2 dy = \int_0^h Y(z)\left[f(z)+b'\int_0^z \rho(\xi)\,d\xi\right]^2 dz, \quad (6.75)$$

where Y is the biaxial modulus, $Y = C_{11} + C_{12} - 2C_{12}^2/C_{11}$, and C_{11} and C_{12} are the elastic stiffness constants. The dislocation density ρ is always positive, but b' may be positive or negative, depending on the sign of the mismatch strain that is being relieved. The line energy of dislocations per unit area, assuming two orthogonal networks with equal cross-sectional density, is

$$E_d = 2\int_0^h F_d(z)\rho(z)dz, \tag{6.76}$$

where

$$F_d(z) \approx \frac{G(z)b^2(z)\left[1-\nu(z)\cos^2\alpha\right]}{4\pi\left(1-\nu(z)\right)}\left[\ln\left(\frac{h-z}{b(z)}\right)+1\right]. \tag{6.77}$$

The equilibrium configuration may be found by minimizing $E_e + E_d$, and this is the generalized semiconductor heterostructure energy minimization approach described by Bertoli et al.[4] and summarized in Section 6.3. Using this approach, Sidoti et al. found the critical layer thickness for sublinear (convex-up) exponentially graded layers of $In_xGa_{1-x}As$ on GaAs (001) with $x(z) = x_\infty(1-e^{-z/\gamma})$, where the limiting indium mole fraction x_∞ was set to 0.1, 0.2, and 0.3. The thickness parameter γ was varied from 0.1 to 1.0 μm. The assumed material parameters are given in Table 6.3; for energy minimization calculations, linear interpolation was used to estimate the lattice and elastic constants for the ternary alloy $In_xGa_{1-x}As$. In the case of the force balance model, b and ν were estimated for the composition $x_\infty/2$.

Figure 6.29 shows the critical layer thickness as a function of the parameter γ for exponential layers with $x_\infty = 0.1$, 0.2, and 0.3. The critical layer thickness increases monotonically with the thickness parameter γ and decreases with x_∞. Figure 6.30 illustrates exponential profiles in composition $x = 0.2[1 - \exp(-z/\gamma)]$, with $x_\infty = 0.2$ and with $\gamma = 0.1$, 0.2, and 0.4. Also indicated in this figure are the critical layer thicknesses for these three cases.

Note that the critical layer thicknesses presented here are for thermal equilibrium, and do not take into account kinetic limitations associated with the nucleation, glide, and multiplication of dislocations. For growth on substrates with high crystalline perfection, such as Si, the number of grown-in dislocations is insufficient to give rise to the observed lattice relaxation. Therefore, dislocation nucleation and multiplication are important and the measured thickness for observable lattice relaxation may be considerably greater than the equilibrium value. Nonetheless, the equilibrium critical layer thickness is the starting point for development of a kinetic model for lattice relaxation.

The simple force balance calculation predicts a smaller value of the critical layer thickness than the energy minimization calculation. As pointed out above, the force balance calculation is based on the assumption that the first misfit dislocation is introduced at the interface with the substrate ($z = 0$). If misfit dislocations are instead introduced at a distance z_d from the interface, then the glide force acting on the grown-in dislocation is reduced for a given thickness of deposit. It will therefore be necessary to grow a thicker

TABLE 6.3

Material Properties Used for InAs, GaAs, and the Ternary Alloy $In_xGa_{1-x}As$

Parameter	InAs	$In_xGa_{1-x}As$	GaAs
a (nm)	0.60584	0.56534 + x(0.0405)	0.56534
b (nm)	0.428	0.400 + x(0.028)	0.400
C_{11} (GPa)	83.3	118.4 – x(35.1)	118.4
C_{12} (GPa)	45.3	53.7 – x(8.4)	53.7
α	60°	60°	60°
λ	60°	60°	60°

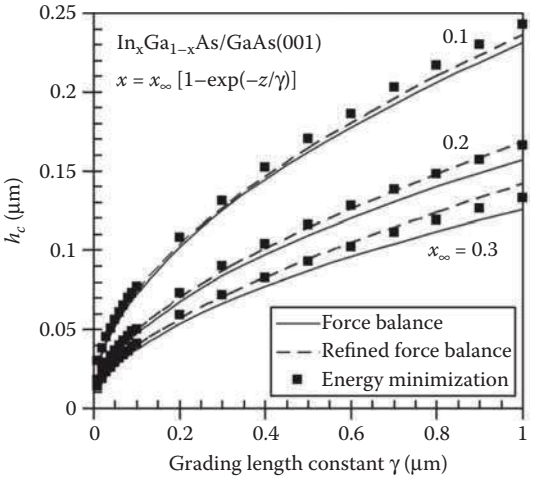


FIGURE 6.29
Critical layer thickness h_c as a function of the grading length constant with x_∞ as a parameter, for (convex-up) exponentially graded $In_xGa_{1-x}As/GaAs(001)$ heteroepitaxial layers as determined by a simple force balance model (solid curve), a refined force balance model taking into account z_c (dashed curves), and energy minimization (solid squares). (Reprinted from D. Sidoti et al., *J. Electron. Mater.*, 29, 1140 [2010]. With permission. Copyright 2010, Minerals, Metals and Materials Society.)

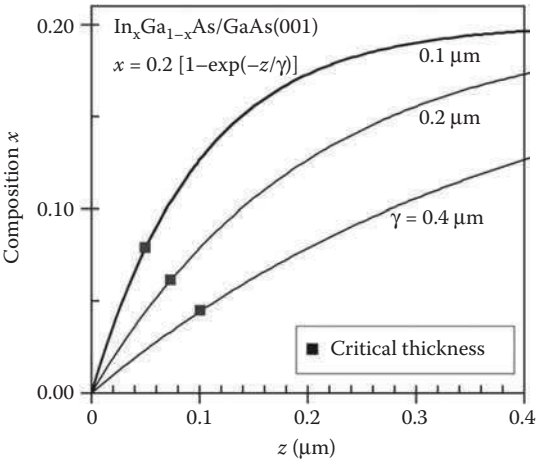


FIGURE 6.30
Exponential compositional profiles for $In_xGa_{1-x}As$ graded layers with $x = 0.2[1-\exp(-z/\gamma)]$, in which $x_\infty = 0.2$ and $\gamma = 0.1$, 0.2, and 0.4. Also indicated in this figure are the critical layer thicknesses for these three cases. (Reprinted from D. Sidoti et al., *J. Electron. Mater.*, 29, 1140 [2010]. With permission. Copyright 2010, Minerals, Metals, and Materials Society.)

epitaxial layer before the glide force is sufficient to cause the grown-in dislocation to glide and create a length of misfit dislocation; that is, the critical layer thickness will be increased. To investigate this, we determined the distance from the interface z_d at which misfit dislocations are first introduced using energy minimization calculations.

Exponentially graded layers with $x_\infty = 0.1$ and $\gamma = 0.1$ μm were first considered. Figure 6.31 shows the equilibrium misfit dislocation density distribution for layers with thicknesses of

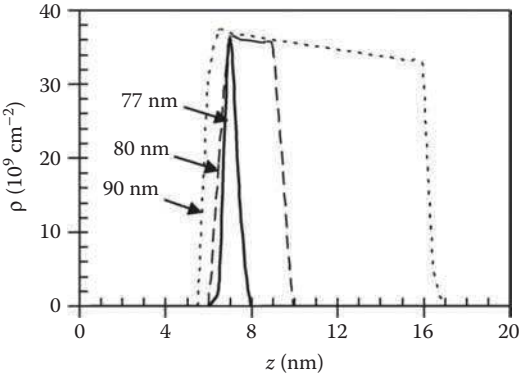


FIGURE 6.31
Misfit dislocation density as a function of distance from the interface for exponentially graded $In_xGa_{1-x}As$ layers/GaAs (001) layers of various thicknesses. The indium composition is assumed to have an exponential profile $x = x_\infty(1 - e^{-z/\gamma})$, where $x_\infty = 0.1$ and $\gamma = 0.1$ μm. This results in a critical layer thickness $h_c \approx 77$ nm. (Reprinted from D. Sidoti et al., *J. Electron. Mater.*, 29, 1140 [2010]. With permission. Copyright 2010, Minerals, Metals and Materials Society.)

77, 80, and 90 nm. Here, 77 nm is the minimum thickness for which energy minimization calculations predict the presence of misfit dislocations in the structure, so this is considered to be the critical layer thickness. The first misfit dislocations are introduced at a distance $z_c \approx 7$ nm from the interface. As the film thickness is increased, the width of the dislocated region increases. This is because the misfit dislocation density is limited by the grading in the layer, while the integrated misfit dislocation density must increase as the layer relaxes. Therefore, dislocations are eventually introduced at distances less than z_c from the interface. In consideration of the critical layer thickness, however, we only need to consider the point z_c where dislocations are *first* introduced.

There are two reasons for this phenomenon. First, the lattice mismatch strain approaches zero at the interface and the introduction of misfit dislocations in this region would actually increase the overall strain energy. Second, misfit dislocations closer to the free surface have reduced line energy.

Figure 6.32 shows z_c as a function of the grading length constant γ, with x_∞ as a parameter. The distance from the interface at which misfit dislocations are first introduced increases monotonically with the grading constant and decreases with the limiting indium mole fraction, and may be estimated using the approximation

$$z_c \approx 18\,\text{nm}\left(\frac{\gamma/1\mu\text{m}}{x_\infty/0.1}\right)^{0.33}. \tag{6.78}$$

When a grown-in dislocation glides to create a length of misfit dislocation at a distance z_c from the interface, the glide force opposing the line tension is

$$F_{G1} = \frac{2Gb(1+\nu)\cos\lambda}{(1-\nu)}\int_{z_c}^{h} f_\infty\left[1-e^{-z/\gamma}\right]dz. \tag{6.79}$$

The line tension in the misfit segment is reduced if this segment is located above the interface:

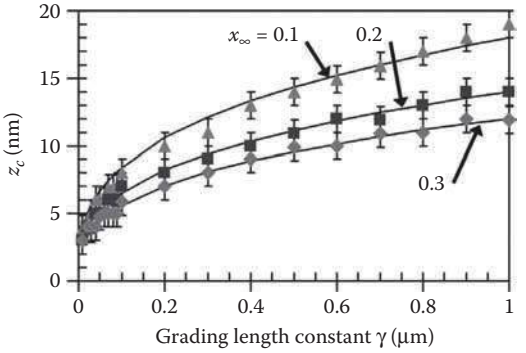


FIGURE 6.32
Distance from interface where the first misfit dislocations are introduced, z_c, as a function of the grading constant γ with x_∞ as a parameter, for (convex-up) exponentially graded $In_xGa_{1-x}As/GaAs$ (001) heteroepitaxial layers. (Reprinted from D. Sidoti et al., *J. Electron. Mater.*, 29, 1140 [2010]. With permission. Copyright 2010, Minerals, Metals and Materials Society.)

$$F_L = \frac{Gb^2\left(1-\nu\cos^2\alpha\right)}{4\pi(1-\nu)}\left[\ln\left(\frac{h-z_c}{b}\right)+1\right]. \tag{6.80}$$

The reduction in the glide force is dominant, because of the logarithmic dependence of the line tension on the cutoff parameter ($h-z_c$). Therefore, the net effect is to increase the critical layer thickness, as predicted by the energy minimization calculations. Taking into account the displacement of the misfit dislocations from the interface, we can make a refined force balance calculation as follows:

$$h_c = \frac{b\left(1-\nu\cos^2\alpha\right)}{8\pi\left|f_\infty\right|(1+\nu)\cos\lambda}\left[\ln\left(\frac{h_c-z_c}{b}\right)+1\right]+z_c+\gamma e^{-z_c/\gamma}-\gamma e^{-h_c/\gamma}. \tag{6.81}$$

Including the effect of z_c with the average values of b and ν,

$$b_{\text{ave}} = b_{GaAs}+\left(b_{InAs}-b_{GaAs}\right)x_\infty\left[1+\frac{\gamma}{h}\left(e^{-h/\gamma}-1\right)\right], \tag{6.82}$$

and

$$\nu_{\text{ave}} = \nu_{GaAs}+\left(\nu_{InAs}-\nu_{GaAs}\right)x_\infty\left[1+\frac{\gamma}{h}\left(e^{-h/\gamma}-1\right)\right]. \tag{6.83}$$

This refined model has been used to calculate the critical layer thickness for comparison to the simple force balance model and the energy minimization calculations, as shown in Figure 6.29. This figure shows that the refined force balance model provides better accuracy than the simple force balance model, though the use of effective values of b and ν introduces some error.

For the design of exponentially graded buffer $In_xGa_{1-x}As$ layers in devices, the critical layer thickness determined by energy minimizations may be conveniently estimated using the approximation

$$h_c \approx 0.243\mu m\left(\gamma/1\mu m\right)^{0.5}\left(x_\infty/0.1\right)^{-0.54}, \tag{6.84}$$

or may be determined using the refined force balance model with less than 5% error over the range of parameters considered by Sidoti et al.[117]

6.8.2 Strain Relaxation in Sublinearly Graded Layers

Salviati et al.[17] have modeled nonlinear buffers having square root and parabolic profiles with the assumption that the strain energy remains constant after the critical thickness is exceeded. This was based on the experimental finding that the residual strain in uniform layers greater than the critical layer thickness may be fit by the relationship

$$\varepsilon^2 h = K, \tag{6.85}$$

where:

- ε is the residual strain
- h is the layer thickness
- K is a fitting parameter[119–121]

To apply the constant strain energy approximation to arbitrarily graded buffers, Salviati et al. assumed composition-independent elastic constants and applied the relationship

$$\int \varepsilon^2 dz = K \tag{6.86}$$

to buffer layers with power-law type profiles of the form $f = Az^\alpha$, where A is a constant. A key assumption was constant lattice relaxation, $\delta = f - \varepsilon =$ constant. The resulting calculations predicted the largest residual surface strain in the superlinear buffer with $\alpha = 2$, the smallest surface strain in the sublinear buffer with $\alpha = 0.5$, and an intermediate value for the linear case with $\alpha = 1$. In an experimental investigation of MBE-grown $In_xGa_{1-x}As$/GaAs (001) heterostructures with sublinear, linear, and superlinear grading of the indium composition, Salviati et al.[17] found by XTEM analysis that the misfit dislocation density scaled with the compositional gradient, and that the structure with sublinear grading had the best quality, as measured by PL and HRXRD. They also found the residual surface strains and MDFZ thicknesses were similar to those predicted by the constant strain energy model, and proposed the use of such a model for the design of graded buffer layers. However, the model does not consider relaxation kinetics and might have limited applicability to low-temperature buffer layers.

6.8.3 Comparison of Sublinearly and Superlinearly Graded Layers

Choi et al.[77] compared sublinear, linear, and superlinear $In_xAl_{1-x}As$ buffers for the MBE growth of $In_xGa_{1-x}As/In_yAl_{1-y}As$ quantum wells on GaAs (001) substrates. The indium mole fraction profiles were of the form

$$x = x_0 + (x_\infty - x_0)(z/h)^\alpha, \tag{6.87}$$

resulting in sublinear grading for $\alpha < 1$, superlinear grading for $\alpha > 1$, and linear grading for $\alpha = 1$. It was found that use of a sublinear buffer with $\alpha = 0.5$ resulted in the highest PL intensity and lowest threading dislocation density of the three profiles. The superlinear profile with $\alpha = 2$ exhibited the lowest PL intensity and highest threading dislocation density, while the sample with linear grading was between the two other cases. These differences suggest the importance of the MDFZ adjacent to the device; the sublinear profile would be expected to have the thickest MDFZ, while the superlinear profile would have the least MDFZ thickness.

Kujofsa and Ayers[122] considered sublinear and superlinear structures by employing a logarithmic-based compositional profile in $In_xGa_{1-x}As$ and $GaAs_{1-y}P_y$ metamorphic buffer layers (MBLs) grown on GaAs (001) substrates. For both of these material systems, they studied the evolution of the equilibrium lattice relaxation and showed that differences in the elastic stiffness coefficients between $In_xGa_{1-x}As$ and $GaAs_{1-y}P_y$ structures with exact lattice mismatch profiles give rise to differences in the equilibrium configuration. The lattice mismatch profile in the nonlinear logarithmic-graded MBL was considered to be

$$f(z) = f_0 + \left(f_h - f_0\right)\left[\frac{\ln\left(1+\frac{z}{h}\right)}{\ln(2)}\right]^m, \tag{6.88}$$

where:

- f_0 is the lattice mismatch at the substrate–MBL interface
- f_h is the value of lattice mismatch at the top of the MBL with thickness h
- z is the distance from the interface
- m is the power grading coefficient
- h is the epilayer thickness

In this work, they considered heterostructures with power grading coefficients of $m = 1$ and $m = 2$. For $m = 1$, the logarithmically graded MBL has a convex-up compositional profile (sublinear), whereas for heterostructures with $m = 2$, the concavity of the compositional profile changes sign to convex down (superlinear). The equilibrium configuration was determined by energy minimization.[4] Based on these results, Kujofsa and Ayers[123] developed analytical models for the in-plane strain and misfit dislocation density in these logarithmically graded MBLs and applied them to the $In_xGa_{1-x}As$-on-GaAs (001) material system. In addition, they also solved for the widths of the interfacial and surface MDFZs. For the determination of the misfit dislocation density, it was considered that the edges of the dislocated region were at distances of z_1 and z_2 from the substrate interface, and that the misfit dislocation density in this dislocated region ($z_1 \le z \le z_2$) was just sufficient to relax the strain associated with the compositional grading. Then in the dislocated region,

$$\rho(z) = \frac{1}{b'}\left|\frac{df(z)}{dz}\right|, \tag{6.89}$$

where b' is the misfit-relieving component of the Burgers vector in the plane of the interface. Therefore, the approximate misfit dislocation density profile is given by

$$\rho = \begin{cases} 0, & z < z_1; \\ \dfrac{m(f_h - f_0)}{b'(h+z)} \dfrac{\ln\left(1+\dfrac{z}{h}\right)^{m-1}}{\ln(2)^m}, & z_1 \le z \le z_2; \quad \text{and} \\ 0, & z > z_2. \end{cases} \tag{6.90}$$

With the approximate misfit dislocation density described above, the depth profile of in-plane strain is given by

$$\varepsilon_{\parallel} = \begin{cases} f_0 + (f_h - f_0) \dfrac{\ln\left(1+\dfrac{z}{h}\right)^m}{\ln(2)^m}, & z < z_1; \\ f_0 + (f_h - f_0) \dfrac{\ln\left(1+\dfrac{z_1}{h}\right)^m}{\ln(2)^m}, & z_1 \le z \le z_2; \quad \text{and} \\ f_0 + (f_h - f_0) \left[\dfrac{\ln\left(1+\dfrac{z}{h}\right)^m}{\ln(2)^m} + \dfrac{\ln\left(1+\dfrac{z_1}{h}\right)^m}{\ln(2)^m} - \dfrac{\ln\left(1+\dfrac{z_2}{h}\right)^m}{\ln(2)^m} \right], & z > z_2. \end{cases} \tag{6.91}$$

An interesting aspect of compositionally graded layers is that they provide a large built-in strain at the surface of the epilayer and allow flexibility in controlling the location of the dislocated region. Therefore, the proper design of sublinear and superlinear metamorphic structures requires an understanding of the important characteristics, such as the widths of the MDFZs, the misfit dislocation density, and the built-in strain. For a given profile of lattice mismatch, the choice of material system will also play a role in the relaxation process by way of the elastic constants and kinetic factors. In their work, Kujofsa and Ayers investigated the strain relaxation and misfit dislocations for log-graded InGaAs/GaAs and GaAsP/GaAs for the case of equilibrium.

Figure 6.33 compares the average equilibrium misfit dislocation density for 250 nm thick layers of logarithmic-graded $In_xGa_{1-x}As/GaAs(001)$ and $GaAs_{1-y}P_y/GaAs$ (001) as a function of the ending lattice mismatch. Sublinear ($m = 1$) and superlinear ($m = 2$) cases are shown. The misfit dislocation density scales in approximate linear fashion with ending mismatch, and the use of a superlinear grading profile gives rise to lower average mismatch and therefore lower average dislocation densities. At low ending mismatch, there is a small difference of about ~5×10^8 cm^{-2} between the use of a sublinear and superlinear compositional profile; however, at higher ending mismatch (~0.85%), the difference becomes more apparent and results in a 2×10^9 cm^{-2} difference between the two types of compositional profiles. In addition, structures with $GaAs_{1-y}P_y$ as the epilayer material contain slightly higher misfit dislocation densities. At an ending mismatch of 1.35%, structures with $GaAs_{1-y}P_y$ as the epilayer material contain 1.5×10^9 cm^{-2} more misfit dislocations than InGaAs, and this is a consequence of the difference in the elastic stiffness constants.

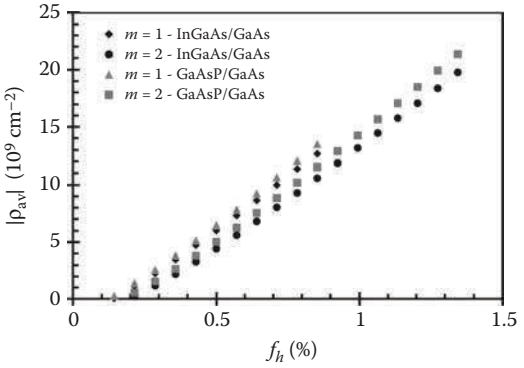


FIGURE 6.33
Average equilibrium misfit dislocation density for 250 nm thick log-graded layers of $In_xGa_{1-x}As/GaAs(001)$ and $GaAs_{1-y}P_y/GaAs(001)$ as a function of the ending lattice mismatch. The case of $m = 1$ corresponds to sublinear grading, while $m = 2$ corresponds to superlinear grading. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Vac. Sci. Technol. B*, 32, 031205 [2014]. With permission. Copyright 2014, American Vacuum Society.)

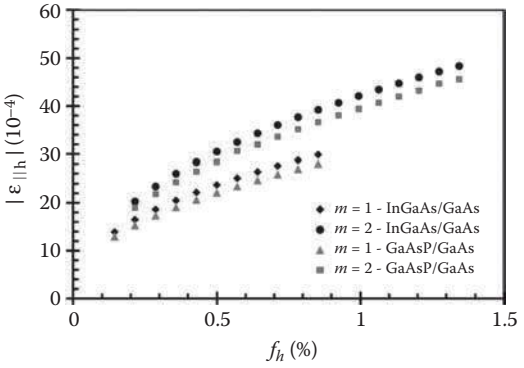


FIGURE 6.34
Surface in-plane strain for 250 nm thick log-graded layers of $In_xGa_{1-x}As/GaAs(001)$ and $GaAs_{1-y}P_y/GaAs(001)$ as a function of the ending lattice mismatch. The case of $m = 1$ corresponds to sublinear grading, while $m = 2$ corresponds to superlinear grading. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Vac. Sci. Technol. B*, 32, 031205 [2014]. With permission. Copyright 2014, American Vacuum Society.)

Figure 6.34 compares the surface in-plane strain for the same structures considered in Figure 6.33. The surface in-plane strain exhibits sublinear behavior with increasing ending mismatch, and heterostructures with a sublinear grading profile contain a higher residual strain due to lower grading rate at the surface of the epilayer. The use of $In_xGa_{1-x}As$ as the material system yields slightly higher strain values than $GaAs_{1-y}P_y$, with the difference becoming more apparent at higher overall lattice mismatch. This was attributed to the higher elastic stiffness constants in $GaAs_{1-y}P_y$; with approximately fixed strain energy per unit area, a material with larger elastic constants will exhibit lower strain.

In terms of device applications, the sublinear profile provides a thicker MDFZ adjacent to the device (see Figure 6.35), which may be beneficial in sweeping out the arms of threading dislocations, whereas the superlinear profile exhibits a thicker MDFZ adjacent to the substrate, and may give rise to fewer pinning interactions during the early stages of growth.

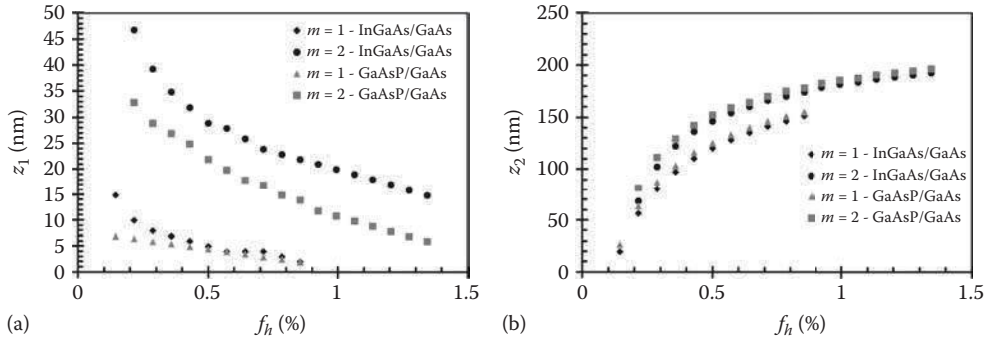


FIGURE 6.35
Thickness of the (a) interfacial and (b) surface MDFZ, for 250 nm thick log-graded layers of $In_xGa_{1-x}As/GaAs(001)$ and $GaAs_{1-y}P_y/GaAs(001)$ as a function of the ending lattice mismatch. The case of $m = 1$ corresponds to sublinear grading, while $m = 2$ corresponds to superlinear grading. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Vac. Sci. Technol. B*, 32, 031205 [2014]. With permission. Copyright 2014, American Vacuum Society.)

6.9 S-Graded Layer

To obtain both of these advantages, by tapering the grading coefficient near the interface and the device layer, Xhurxhi et al.[125] proposed the use of an complementary error function, or S-graded compositional profile. The lattice mismatch profile in an S-graded buffer layer is given by

$$f = \begin{cases} \dfrac{f_h}{2}\left[-\text{erf}\left(\dfrac{\mu - z}{\sigma\sqrt{2}}\right) + \text{erf}\left(\dfrac{\mu}{\sigma\sqrt{2}}\right)\right], & z < \mu; \\ \dfrac{f_h}{2}, & z = \mu, \\ \dfrac{f_h}{2}\left[\text{erf}\left(\dfrac{z - \mu}{\sigma\sqrt{2}}\right) + \text{erf}\left(\dfrac{\mu}{\sigma\sqrt{2}}\right)\right], & z > \mu \end{cases} \tag{6.92}$$

where:

- f_h is the limiting mismatch
- σ is the standard deviation parameter
- μ is the mean parameter
- z is the distance from the interface

The parameters μ, σ, f_h, and h can be chosen by the crystal grower to obtain the desired buffer layer characteristics, and three particular S-graded mismatch profiles are illustrated in Figure 6.36a.

6.9.1 Misfit Dislocations and Strain in the S-Graded Layer

Xhurxhi et al.[124] determined the equilibrium misfit dislocation density and strain profiles for S-graded buffer layers of $In_xGa_{1-x}As$ on GaAs (001) substrates using the generalized semiconductor heterostructure energy minimization technique. These results can be used to design S-graded layers for the achievement of desired peak misfit dislocation density or

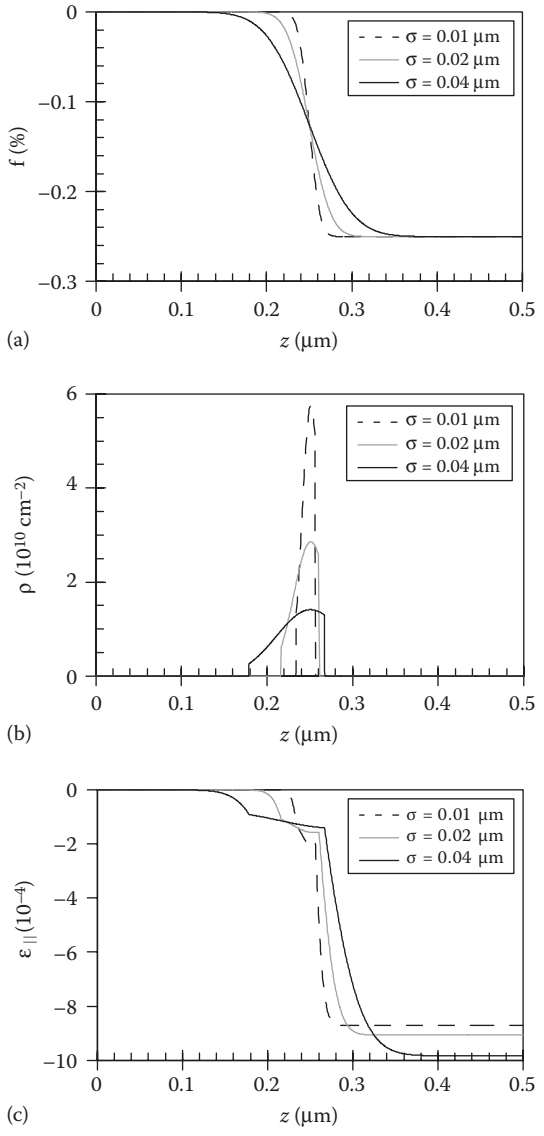


FIGURE 6.36
Characteristics of 0.5 μm thick S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrates with $x_h = 0.035$ (corresponding to $f_h = -0.5\%$) with mean parameter μ = 0.25 μm and standard deviation parameter values of σ = 0.01 μm, 0.02 μm, and 0.04 μm. (a) Lattice mismatch, (b) misfit dislocation density, and (c) in-plane strain. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

MDFZ thicknesses, and also serve as the starting point for kinetic calculations, which will enable the determination of the threading dislocation densities in structures grown under nonequilibrium conditions.

Figure 6.36 shows the profiles of lattice mismatch, misfit dislocation density, and in-plane strain in S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrates with 0.5 μm total thickness and standard deviation parameters of σ = 0.1, 0.2, and 0.4 μm. In each case, the initial

composition at the interface was zero (lattice matched to GaAs), while the final composition was $x_h = 0.035$, corresponding to $f_h = -0.5\%$. The mean parameter was fixed at one-half the buffer thickness. The material parameters for GaAs were assumed to be $a = 0.56534$ nm, $C_{11} = 118.4$ GPa, and $C_{12} = 53.7$ GPa; the corresponding values for $In_xGa_{1-x}As$ were assumed to be linear functions of the composition: $a = (0.56534 + x0.405)$nm, $C_{11} = (118.4 - x35.1)$GPa, and $C_{12} = (53.7 - x8.4)$GPa.

Though the S-graded buffer layers exhibit rather complex behavior, three important general characteristics may be observed in the calculated results. First, there are MDFZs at the bottom of the S-graded layer (adjacent to the substrate interface) and the top of the layer (adjacent to the free surface). The edges of these MDFZs are located at z_1 and z_2, so the thicknesses of the bottom and top MDFZs are z_1 and $h - z_2$, respectively. Second, there is a dislocated region with a thickness $z_2 - z_1$ in which the misfit dislocation density takes on a truncated Gaussian profile. Third, there is a nearly uniform in-plane strain in the top MDFZ. The equilibrium strain in this region can be relatively large, and for the cases shown in Figure 6.36, its value is more than twice the equilibrium strain ($\varepsilon_{\parallel} \approx -4 \times 10^{-4}$) for a 0.5 μm thick uniform layer of $In_{0.035}Ga_{0.965}As$ on GaAs (001).

The existence of the interfacial MDFZ is expected on the basis of minimum energy. With zero lattice mismatch at $z = 0$, the introduction of misfit dislocations would increase both the strain energy and dislocation line energy, so misfit dislocations are not expected to form right at the interface. There is a finite distance from the interface where the line energy cost of misfit dislocations is balanced by the strain energy they release in the growing film, and this dictates a finite thickness of misfit dislocation free material near the interface. The thickness of this interfacial MDFZ depends on the details of the lattice mismatch profile, and therefore f_h, σ, and μ.

The formation of the surface MDFZ may also be understood from the point of view of energy minimization. Although this MDFZ exists in material with significant lattice mismatch, a significant portion of the strain is relaxed by defects in the underlying dislocated zone. Because of this, and due to the proximity to the surface, relatively little strain energy can be released by the introduction of misfit dislocations in this near-surface material. At the same time, line energies of misfit dislocations near the surface are not reduced significantly because of the weak logarithmic dependence of the line energy on the distance from the interface. The consequence is that there is a finite thickness of material near the surface in which the introduction of misfit segments is not energetically favorable. In a qualitative sense, this is analogous to the behavior predicted[5] and observed[65] in linearly graded layers.

The shape of the misfit defect density profile in the dislocated region is expected to be a truncated Gaussian if it is assumed that the misfit dislocation density is just sufficient to relax the strain introduced by the compositional grading, as has been found to be the case in linearly graded layers. If b' is the misfit-relieving component of the Burgers vector, then the areal density of misfit dislocations will be $\rho = |df/dz|/b'$. For the S grading compositional profile considered here, $|df/dz|$ is Gaussian in character, but the profile is truncated by the existence of the MDFZs, as explained above.

Heteroepitaxial layers are not deposited in equilibrium, and tend to exhibit fewer misfit dislocations than predicted by equilibrium models. Nonetheless, we can expect S-graded buffer layers to exhibit the general features outlined here, including the dislocated region sandwiched between two MDFZs. These general characteristics of S-graded buffer layers may promote longer misfit segments and improve their performance in threading dislocation reduction as follows. (1) As in the linearly graded layer, new misfit dislocations are introduced at the top of the growing material, where there are relatively few existing

dislocations, and the reduction of the dislocation–dislocation interactions that give rise to pinning or decreased dislocation mobility allow the uninhibited glide of dislocations to form longer misfit segments. Improved performance is expected in the S-graded structure due to the wide interfacial MDFZ and the tapered misfit dislocation density above it. (2) The MDFZ near the surface prevents the introduction of new dislocation loops near the surface in the final stages of buffer layer growth, and instead, existing misfit dislocations are allowed to grow in length by glide. The S-graded buffer can have a thicker surface MDFZ than the linearly graded layer, which is preserved even after the growth of a top device layer. (3) A graded layer has increased residual strain in its surface MDFZ relative to a uniform buffer, which can promote higher effective stresses and misfit dislocation velocities. In the S-graded layer, the average strain in the top MDFZ is even greater than in the linearly graded layer, and that is expected to improve the threading dislocation reduction performance.

Xhurxhi et al.[124] have demonstrated some of these potential advantages of S-graded layers compared with linearly graded or uniform layers using energy minimization calculations, as illustrated in Figure 6.37. Figure 6.37a compares the equilibrium in-plane strain profiles for S-graded (with σ = 0.04 μm), linearly graded, and uniform layers of $In_xGa_{1-x}As$ on GaAs (001), all with thickness 0.5 μm and a top indium composition of $x = 0.035$. Both

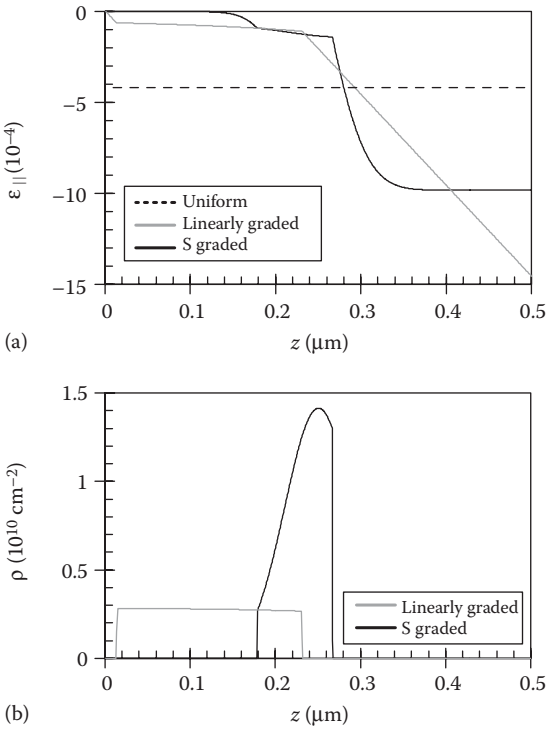


FIGURE 6.37
Comparison of 0.5 μm thick $In_xGa_{1-x}As$ layers on GaAs (001) substrates, all with a top composition of $x_{ti} = 0.035$, but with different compositional profiles. (a) In-plane strain profiles for layers with S grading (σ = 0.04 μm), linear grading, and uniform composition. (b) Misfit dislocation density profiles in layers with S grading and linear grading. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

the linearly graded and S-graded structures exhibit enhanced surface strain compared with the uniform layer. Although the linearly graded layer actually presents the highest strain at its surface, the S-graded case provides a thicker zone of enhanced strain that can provide more benefit in terms of increasing the lengths of misfit dislocations and reducing the threading density. The equilibrium misfit dislocation density profiles for the S-graded and linearly graded buffers are compared in Figure 6.37b. Whereas S grading gives rise to thick MDFZs at the bottom and top of the layer, the linear layer has the undesirable characteristic of a high-misfit-dislocation density near the substrate interface. The S-graded layer shows a higher peak misfit dislocation density than the linearly graded film, but this peak misfit dislocation density may be tailored by the choice of the standard deviation parameter.

In order to facilitate the design of S-graded buffer layers for devices, Xhurxhi et al. developed an approximate model for $In_xGa_{1-x}As$ S-graded layers on GaAs (001) substrates. This model is based on a phenomenological model for the edges of the dislocated region:

$$z_1 \approx \mu - \sigma\left(\frac{|f_h|}{f_{01}}\right)^{n_1} \tag{6.93}$$

and

$$z_2 \approx \mu + \sigma\left(\frac{|f_h|}{f_{02}}\right)^{n_2}, \tag{6.94}$$

where $f_{01} = 0.0035$, $n_1 = 0.2$, $f_{02} = 0.0150$, and $n_2 = 0.75$. These models are compared with numerical results obtained by energy minimization in Figure 6.38. The thickness of the dislocated region, $z_2 - z_1$, is independent of the total thickness as long as $h > 6\sigma$, and scales directly with the standard deviation parameter:

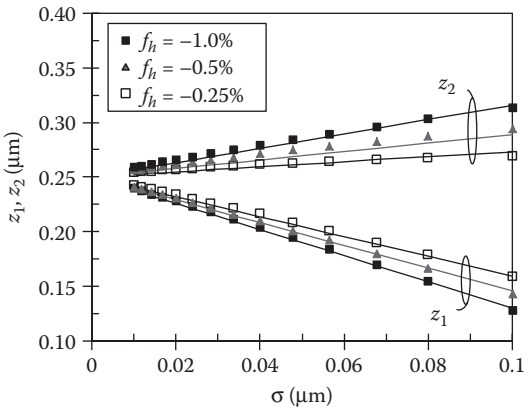


FIGURE 6.38
Boundary points z_1 and z_2 of the dislocated region for 0.5 μm thick S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrates with mean parameter μ = 0.25 μm and top lattice mismatch f = –1.00%, –0.50%, and –0.25%. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

$$z_2 - z_1 \approx \sigma\left[\left(\frac{|f_h|}{f_{01}}\right)^{n_1} + \left(\frac{|f_h|}{f_{02}}\right)^{n_2}\right]. \tag{6.95}$$

Although these are applicable to S-graded layers of $In_xGa_{1-x}As/GaAs$ (001), similar expressions may be developed for other material systems by fitting the numerical energy minimization results. Once the edges of the dislocated region are determined, the strain and misfit dislocations in the S-graded buffer may be modeled as follows.

In the dislocated region ($z_1 \le z \le z_2$), the in-plane strain is approximately constant so the misfit dislocation density is just sufficient to relax the strain due to the grading in the layer:

$$\rho = \frac{1}{b'}\left|\frac{df}{dz}\right|, \tag{6.96}$$

where b' is the misfit-relieving component of the Burgers vector in the plane of the interface. Therefore, the misfit dislocation density profile is

$$\rho = \begin{cases} 0, & z < z_1; \\ \dfrac{f_h}{b'\sigma\sqrt{2\pi}} \exp\left[\dfrac{-(z-\mu)^2}{\sigma^2/2}\right], & z_1 \le z \le z_2; \text{ and} \\ 0, & z > z_2. \end{cases} \tag{6.97}$$

The resulting misfit dislocation density profile is a truncated Gaussian, with the peak misfit dislocation density of

$$\rho_{max} = \frac{f_h}{b'\sigma\sqrt{2\pi}} \tag{6.98}$$

occurring at $z = \mu$. The peak misfit dislocation density scales inversely with the standard deviation parameter, so the design of the S-graded layer involves a trade-off between the width of the dislocated region and the peak dislocation density.

Considering the equilibrium strain, the bottom MDFZ ($z \le z_1$) is coherently strained, so the in-plane strain is equal to the lattice mismatch,

$$\varepsilon_{\parallel} = \frac{f_h}{2}\left[-\text{erf}\left(\frac{\mu - z}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right]; z \le z_1, \tag{6.99}$$

and the strain at $z = z_1$ is

$$\varepsilon_{\parallel}(z_1)\frac{f_h}{2}\left[-\text{erf}\left(\frac{\mu - z_1}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right]. \tag{6.100}$$

For $z_1 \le z \le z_2$ (the dislocated layer) the in-plane strain is approximately constant at the value given above, assuming the misfit dislocation density is just enough to relax

the strain associated with the grading. For $z > z_2$, there are no misfit dislocations, so $d\varepsilon_{\|}/dz = df/dz$. The equilibrium in-plane strain profile in the partially relaxed S-graded layer is thus given by

$$\varepsilon_{\|} = \begin{cases} \left(\frac{f_h}{2}\right)\left[-\text{erf}\left(\frac{\mu - z}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right], & z < z_1; \\ \left(\frac{f_h}{2}\right)\left[-\text{erf}\left(\frac{\mu - z_1}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right], & z_1 \le z \le z_2; \text{ and} \\ \left(\frac{f_h}{2}\right)\left[\text{erf}\left(\frac{z - \mu}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{z_2 - \mu}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{\mu - z_1}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right], & z > z_2. \end{cases} \quad (6.101)$$

Thus, both the in-plane strain and misfit dislocation density profiles in the S-graded buffer layer may be modeled as long as expressions for z_1 and z_2 are known. In practice, it is possible to use either analytical expressions for z_1 and z_2, determined by minimizing the sum of the strain and dislocation line energy with respect to each, or simpler, phenomenological expressions for z_1 and z_2 of the type given above.

Figures 6.39 and 6.40 compare the approximate model with the results of detailed numerical dislocation density calculations, for 0.5 μm thick S-graded $In_xGa_{1-x}As/GaAs$ (001) layers having $f_h = -0.005$ with $\sigma = 0.01$, 0.02, and 0.04 μm. The approximate model predicts the misfit dislocation density with better than 5% accuracy, showing the appropriateness of assuming that the misfit dislocation density in the dislocated region is just sufficient to relieve the strain associated with the grading. The approximate model underestimates the surface strain somewhat, because of the slight variation in the strain in the dislocated region, but provides sufficient accuracy to make it useful in the design of S-graded buffer layers.

Because the intended application is for device buffer layers, Xhurxhi et al. compared the equilibrium dislocation densities in linearly graded and S-graded structures, with and

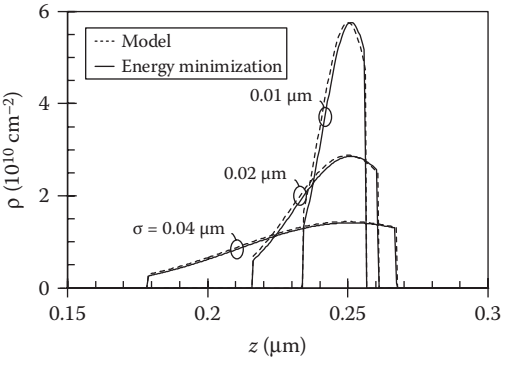


FIGURE 6.39
Misfit dislocation density profiles for 0.5 μm thick S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrates with $x_h = 0.035$ (corresponding to $f_h = -0.5\%$) with mean parameter $\mu = 0.25$ μm and standard deviation parameter values of $\sigma = 0.01$, 0.02, and 0.04 μm. The solid curves were determined by numerical energy minimization calculations, and the dashed curves were obtained using the approximate model. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

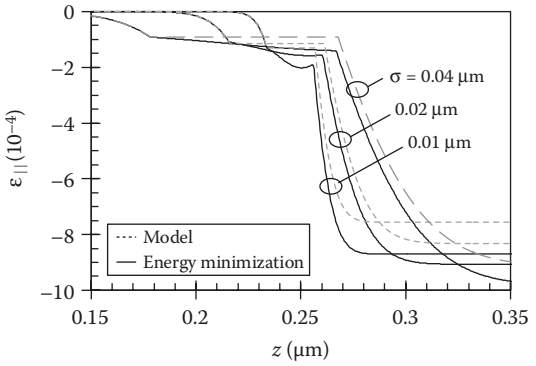


FIGURE 6.40
In-plane strain profiles for 0.5 μm thick S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrates with $x_h = 0.035$ (corresponding to $f_h = -0.5\%$) with mean parameter μ = 0.25 μm and standard deviation parameter values of σ = 0.01, 0.02, and 0.04 μm. The solid curves were determined by numerical energy minimization calculations, and the dashed curves were obtained using the approximate model. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

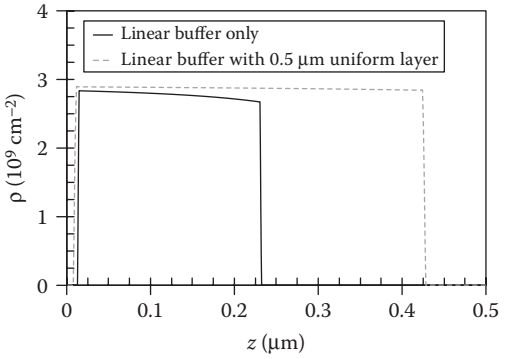


FIGURE 6.41
Behavior of the linearly graded buffer layer with top loading by a uniform device layer. Misfit dislocation density profiles for 0.5 μm thick linearly graded $In_xGa_{1-x}As$ layers on GaAs (001) substrate with $x_h = 0.035$ and no uniform layer on top (solid black curve) and with a 0.5 μm thick uniform layer on top (dashed gray curve). The composition of the uniform layer is $In_{0.035}Ga_{0.965}As$. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

without a uniform device layer deposited on top. Figure 6.41 illustrates the misfit dislocation density in a 0.5 μm thick linearly graded layer of $In_xGa_{1-x}As$, with the indium mole fraction graded from 0% to 3.5%, both with and without a uniform 0.5 μm thick layer of $In_{0.035}Ga_{0.965}As$ on top. The addition of the uniform layer (device layer) on top of the buffer increases the width of the dislocated region from 218 to 418 nm—approximately doubling the dislocated thickness. The top MDFZ in the buffer layer is nearly eliminated because the increase of the dislocated thickness is mostly in the top MDFZ. The behavior is quite different in the case of the S-graded buffer, shown in Figure 6.42, where a ~0.2 μm top MDFZ remains even after the growth of a 0.5 μm uniform top layer. Therefore, the S-graded buffer provides superior performance compared with the linear buffer in terms of maintaining a wide MDFZ adjacent to the device layer.

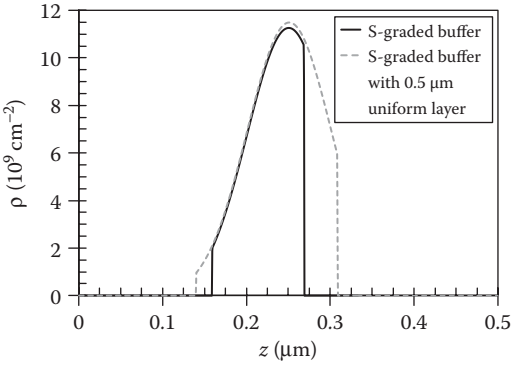


FIGURE 6.42
Behavior of the S-graded buffer layer with top loading by a uniform device layer. Misfit dislocation density profiles for 0.5 μm thick S-graded $In_xGa_{1-x}As$ layers on GaAs (001) substrate with $x_h = 0.035$ and no uniform layer on top (solid black curve) and with a 0.5 μm thick uniform layer on top (dashed gray curve). The composition of the uniform layer is $In_{0.035}Ga_{0.965}As$. (Reprinted from S. Xhurxhi et al., *J. Electron. Mater.*, 40, 2348 [2011]. With permission. Copyright 2011, Minerals, Metals and Materials Society.)

In summary, the S-graded layer exhibits misfit dislocation free regions (MDFZs) near the substrate interface and the free surface (or device interface). The standard deviation parameter may be selected to achieve a particular peak misfit dislocation density, while the mean parameter and total buffer thickness may be designed to achieve the desired MDFZ thicknesses adjacent to the substrate and device layer interfaces. The S-graded buffer layer exhibits two potential advantages compared with the linearly graded buffer layers often used in metamorphic device layers. First, the ability to design the peak misfit dislocation density and MDFZ thicknesses is expected to enable minimization of the density of electronically active threading dislocations at the top surface, and second, the dislocation density profile in the S-graded layer is less susceptible to loading by the growth of a uniform device layer on top, and avoids the penetration of a dislocated region into the device layer itself.

6.9.2 Refined Model for S-Graded Layers

The model for S-graded layers developed by Xhurxhi et al.[124] underestimated the in-plane strain in both the dislocated region and the surface MDFZ. Xhurxhi et al. assumed that the misfit dislocation density was sufficient to relax the mismatch associated with grading, and as such, the in-plane strain was assumed to be constant in this region. Minimum energy calculations show that the strain in the dislocated region actually varies in an approximate linear manner. Kujofsa et al.[125] developed a refined model for the S-graded layer in which the strain is considered to increase linearly within the dislocated region. For its implementation, a factor of $\mp(z/z_1)$ is introduced, as can be seen in the middle expression of Equation 6.102. Physically, this comes about because of the dislocation line energy dependence on distance from the free surface. It should be noted that $\mp$ refers to lattice mismatch being accommodated in tensile or compressive situations, respectively. For $z > z_2$, there are no misfit dislocations, so $d\varepsilon_{\parallel}/dz = df/dz$. Therefore, according to the refined model, the equilibrium strain profile in the S-graded layer is given by

$$\varepsilon_{\parallel} = \begin{cases} \left(\frac{f_h}{2}\right)\left[-\text{erf}\left(\frac{\mu - z}{\sigma\sqrt{2}}\right) + \text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right], & z < z_1; \\ \left(\frac{f_h}{2}\right)\left[-\text{sign}(f)\frac{z}{z_1}\text{erf}\left(\frac{\mu - z_1}{\sigma\sqrt{2}}\right) + \frac{z}{z_1}\text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right], & z_1 \le z \le z_2; \text{ and} \\ \left(\frac{f_h}{2}\right)\begin{bmatrix} \text{erf}\left(\frac{z - \mu}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{z_2 - \mu}{\sigma\sqrt{2}}\right) \\ -\text{sign}(f)\frac{z_2}{z_1}\text{erf}\left(\frac{\mu - z_1}{\sigma\sqrt{2}}\right) + \frac{z_2}{z_1}\text{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \end{bmatrix}, & z > z_2. \end{cases} \tag{6.102}$$

This model was applied to S-graded ZnS_ySe_{1-y}/GaAs (001) heterostructures and was shown to provide a more accurate description of the results obtained by a detailed minimum energy algorithm.

Kujofsa et al.[125] also provided analytical solutions for the equilibrium widths of the interfacial and surface MDFZs by simultaneously solving the sum of the partial derivatives,

$$\frac{\partial\left(E_d + E_e\right)}{\partial z_1} = 0 \tag{6.103}$$

and

$$\frac{\partial\left(E_d + E_e\right)}{\partial z_2} = 0\,, \tag{6.104}$$

where E_d and E_e are the dislocation and strain energies per unit area, respectively, given by

$$E_d = 2\int_{z_1}^{z_2} F_d(z)\rho(z)\,dz \tag{6.105}$$

and

$$E_e = Y\int_0^h \varepsilon_{\parallel}^2(z)\,dz\,. \tag{6.106}$$

The reader is referred to the original reference for this treatment.

6.9.3 Threading Dislocations in S-Graded Layers

Kujofsa and Ayers[126] investigated threading dislocations in ZnS_xSe_{1-x} on GaAs (001) employing an S-graded compositional profile as developed by Xhurxhi et al.[124] and Kujofsa et al.[125] They considered heterostructures where the mean parameter was fixed at $\mu = h/2$, and the standard deviation parameter varied from 20 to 80 nm. Furthermore, the starting sulfur composition in these layers was fixed at 6% (lattice matched to the substrate) and the

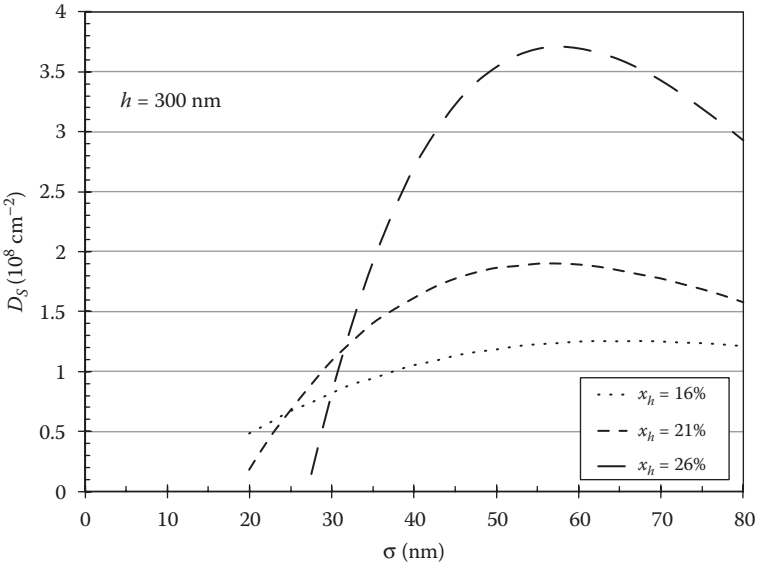


FIGURE 6.43
Surface threading dislocation for 300 nm thick S-graded ZnS_xSe_{1-x}/GaAs (001) layers as a function of the standard deviation parameter. The maximum sulfur composition (x_h) was equal to 10%, 16%, and 21%, respectively. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 23, 1420005 [2014]. With permission. Copyright 2014, World Scientific.)

ending sulfur composition varied: x_h = 16%, 21% and 26%. The thickness (h) of the epilayer ranged from 0.1 to 1.6 μm. For each structure, they studied how the surface threading dislocation behavior depended on the alterations of each individual parameter.

Figure 6.43 shows the surface threading dislocation (D_s) density for 300 nm thick ZnS_xSe_{1-x} layers deposited on GaAs (001) as a function of σ with ending sulfur composition as a parameter. The assumed growth temperature for these layers was 360°C. The results of Figure 6.43 show that the surface threading dislocation exhibits a two-regime behavior with an increasing standard deviation parameter; it can be seen that there exists a critical standard deviation parameter where D_s changes from monotonically increasing to monotonically decreasing for higher σ. In other words, for each case of the top composition, the surface threading dislocation density reaches a maximum and then decreases. In addition, an increase in the ending composition results in a higher threading dislocation density at the critical standard deviation parameter. These results indicate that low threading dislocations (~5×10^7 cm^{-2}) are possible with σ ≤ 30 nm; however, the use of a fixed buffer layer thickness does not afford the necessary flexibility to achieve desired surface threading dislocation density (~10^6 cm^{-2}) for device application. Therefore, in order to achieve higher flexibility in the device design, it will be necessary to tailor both h and σ in order to control the surface dislocation density.

Figure 6.44 illustrates the thickness dependence of the surface threading dislocation and sign of Burgers vector product (sign(b) · D_s) with a varying standard deviation parameter. In contrast to the results of Figure 6.43, it can be seen here that a variation in the buffer layer thickness for a given mismatch and standard deviation parameter may cause the mobile threading dislocation density to vanish at the buffer layer surface. Each type of buffer in Figure 6.44 exhibits three regimes of threading dislocation behavior. In regime

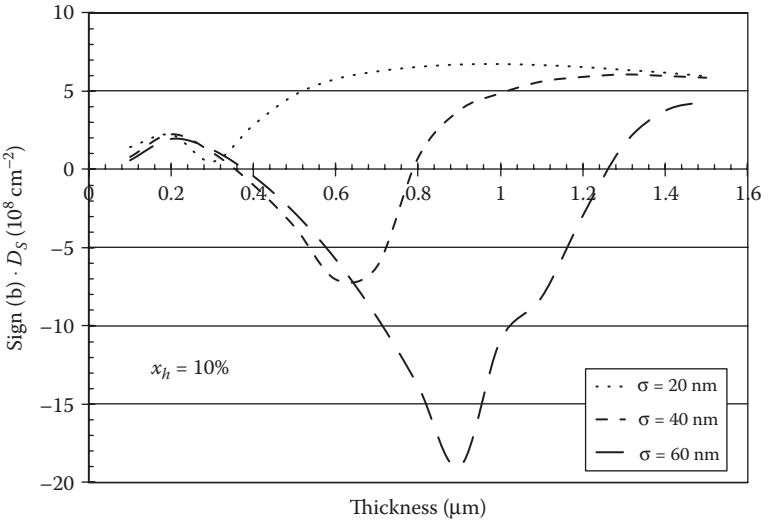


FIGURE 6.44
Thickness dependence of the surface threading dislocation and sign of Burgers vector product (sign(b) · D_s) for an S-graded ZnS_xSe_{1-x}/GaAs (001) layer with a maximum sulfur composition x_h equal to 10% and σ as a parameter. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 23, 1420005 [2014]. With permission. Copyright 2014, World Scientific.)

1, dislocations with positive Burgers vector are introduced to relax compressive mismatch provided in the epilayer; this occurs because of the sizable difference in the thermal expansion coefficients between ZnS_xSe_{1-x} and GaAs, which renders compressive strain in material with 6% sulfur at the growth temperature.

In regime 2, the sulfur composition is sufficient to produce tensile strain, so that dislocations with negative Burgers vector (opposite sense to the lattice mismatch that is relaxing) are introduced. The value of sign(b) · D becomes less positive and may even change sign. In regime 3, as the thickness of the epilayer is increased, because of the thermal mismatch compounded by an increase of annihilation and coalescence reactions, compressive strain requires the introduction of dislocations with positive Burgers vector. The value of sign(b) · D will become more positive, and there can be a second zero crossing, which can be employed to achieve a surface density of zero mobile dislocations.

This work has fascinating implications for applications to metamorphic devices, but additional work will be necessary to fully understand the behavior of S-graded layers.

6.10 Strained Layer Superlattices

Superlattice buffers have been used in conjunction with uniform or graded buffer layers to reduce the surface threading dislocation density and improve the surface smoothness of metamorphic material. There are two basic mechanisms for threading dislocation density reduction by superlattice buffers: dislocation filtering occurs when a superlattice is placed on top of high dislocation density material, and dislocation inhibition occurs in the case

of a superlattice placed below a graded buffer to reduce the steady-state density. In either case, surface smoothness can be improved by suppressing strain-induced cross-hatch.

The dislocation filtering mechanism has been studied to a great extent in the context of GaAs-on-Si.[127] Though GaAs-on-Si typically contains 10^8 to 10^9 cm^{-2} threading dislocations, this density can be reduced significantly by inserting a strained layer superlattice (SLS). The periodic reversals of strain in the SLS bend dislocations back and forth as they pass through, and those with cancelling components of the Burgers vector are bent in opposite directions, thereby promoting annihilation and coalescence reactions.[33] Two 60° dislocations with opposite Burgers vectors can participate in an annihilation reaction, or two 60° dislocations with cancelling tilt components can coalesce to form a single edge dislocation. Depending on the dislocation sources present, there will always be some net Burgers vector content that cannot be eliminated, so perfect filtering has not been achieved.

Yamaguchi et al.[128] demonstrated remarkable reduction of the threading dislocation density in GaAs-on-Si (001) by a combination of thermal cycle annealing and strained layer superlattice dislocation filters. By an experimental and modeling study of $In_xGa_{1-x}As/GaAs$ and $In_xGa_{1-x}As/GaAs_{1-y}P_y$ strained layer superlattices, they clarified the requirements on the SLS thickness for optimum dislocation filtering, and showed that the optimum thickness is less than the Matthews and Blakeslee[1] critical layer thickness for a uniform layer with the same average mismatch.

The dislocation inhibition mechanism of an SLS comes into play when a superlattice is inserted *below* a graded buffer, and has been demonstrated for both step-graded[33] and linearly graded buffers.[73,78] In this instance, the bending back and forth of dislocations may enhance the length of misfit segments, by bending dislocations away from pinning defects. This enables the establishment of a lower steady-state dislocation density in the graded buffer above.

The inclusion of superlattice buffers can also improve the surface smoothness in device structures.[78] This benefit also arises from the periodic strain reversals in the growth direction, which undermine the small local lateral strain fluctuations responsible for roughening and cross-hatch.

6.11 Conclusion

A wide variety of semiconductor devices, including HEMTs, heterojunction bipolar transistors (HBTs), LEDs, and lasers, may be grown on lattice-mismatched substrates using graded and multilayered metamorphic structures. Common challenges in their design include the high densities of threading dislocations, strain-induced surface roughening and cross-hatch, crystallographic tilting, and three-dimensional nucleation. Compositional grading can be tailored to distribute the misfit dislocations and therefore control the threading dislocation density. Low-temperature growth of these graded buffers can improve the surface smoothness and reduce the propagation of threading dislocations to the device layers.

PROBLEMS

1. Suppose $Si_{1-x}Ge_x$ is grown on a Si (001) substrate with a linear profile in grading. The thickness of the graded layer is 0.5 μm and the germanium mole fraction is

graded from 0 to 0.2. Determine the critical layer thickness for the onset of lattice relaxation. Make reasonable assumptions, but state them clearly.

2. Suppose $Si_{1-x}Ge_x$ is grown on a Si (001) substrate with a linear profile in grading. The thickness of the graded layer is 0.5 μm and the germanium mole fraction is graded from 0 to 0.2. Find (a) the approximate thickness of the dislocated region and therefore the MDFZ, (b) the areal density of misfit dislocations in this region, (c) the surface strain in the graded layer, and (d) the average strain in the graded layer. Make reasonable assumptions, but state them clearly.
3. Consider a linearly graded layer of $In_xGa_{1-x}As$ on GaAs (001). The thickness is 0.5 μm. (a) Find the distance from the interface where misfit dislocations are first introduced, assuming the layer is graded to a final indium content of 0.2. (b) Repeat for a final indium content of 0.5.
4. Suppose $GaAs_{1-y}P_y$ is grown on GaAs (001) with a linear profile in composition and a final phosphorus composition of 0.5. Estimate the threading dislocation density if the buffer layer thickness is (a) 0.1 μm, (b) 1 μm, and (c) 10 μm.
5. Consider exponentially graded $In_xGa_{1-x}As$ on GaAs (001) with a total thickness of 0.5 μm and the profile $x = x_\infty(1 - e^{-z/\gamma})$, where $x_\infty = 0.3$ and $\gamma = 0.1$ μm. Find (a) the distance from the interface where dislocations are first introduced, and (b) the critical layer thickness.
6. For S-graded $In_xGa_{1-x}As$ on GaAs (001) with a total thickness of 0.5 μm, a final composition of 0.4, a mean parameter of 0.25 μm, and a standard deviation parameter of 0.1 μm, find the width of the dislocated region and the widths of the two MDFZs.

References

1. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers. I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).
2. D. Sidoti, S. Xhurxhi, T. Kujofsa, S. Cheruku, J. P. Correa, B. Bertoli, P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Initial misfit dislocations in a graded heteroepitaxial layer, *J. Appl. Phys.*, 109, 023510 (2011).
3. B. Bertoli, D. Sidoti, S. Xhurxhi, T. Kujofsa, S. Cheruku, J. P. Correa, P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Equilibrium strain and dislocation density in exponentially graded $Si_{1-x}Ge_x$/Si (001), *J. Appl. Phys.*, 108, 1 (2010).
4. B. Bertoli, E. N. Suarez, J. E. Ayers, and F. C. Jain, Misfit dislocation density and strain relaxation in graded semiconductor heterostructures with arbitrary composition profiles, *J. Appl. Phys.*, 106, 073519 (2009).
5. J. Tersoff, Dislocations and strain relief in compositionally graded layers, *Appl. Phys. Lett.*, 62, 693 (1993); *Appl. Phys. Lett.*, 64, 2748 (1994).
6. T. Kujofsa, W. Yu, S. Cheruku, B. Outlaw, S. Xhurxhi, F. Obst, D. Sidoti, et al., Plastic flow and dislocation compensation in ZnS_ySe_{1-y}/GaAs (001) heterostructures, *J. Electron. Mater.*, 41, 2993 (2013).
7. H. Alexander and P. Haasen, Dislocations and plastic flow in the diamond structure, in *Solid State Physics*, Vol. 22: Advances in Research and Applications (New York: Academic, 1968).
8. B. W. Dodson and J. Y. Tsao, Relaxation of strained-layer semiconductor structures via plastic flow, *Appl. Phys. Lett.*, 51, 1325 (1987); *Appl. Phys. Lett.*, 52, 852 (1988).

9. S. Cheruku, T. Kujofsa, S. Xhurxhi, D. Sidoti, J. P. Correa, B. Bertoli, E. N. Suarez, F. C. Jain, and J. E. Ayers, Plastic flow in semiconductor heterostructures, presented at Connecticut Microlectronics and Optoelectronics Symposium, Storrs, CT (April 7, 2010).
10. S. K. Ghandhi, VLSI *Fabrication Principles: Silicon and Gallium Arsenide*, 2nd ed. (New York: Wiley, 1994).
11. J. W. Matthews, Defects associated with the accommodation of misfit between crystals, *J. Vac. Sci. Technol.*, 12, 126 (1975).
12. M. Tachikawa and M. Yamaguchi, Film thickness dependence of dislocation density reduction in GaAs-on-Si substrates, *Appl. Phys. Lett.*, 56, 484 (1990).
13. A. E. Romanov, W. Pompe, G. E. Beltz, and J. S. Speck, An approach to threading dislocation reaction kinetics, *Appl. Phys. Lett.*, 69, 3342 (1996).
14. S. Akram, H. Ehsani, and I. B. Bhat, The effect of GaAs surface stabilization on the properties of ZnSe grown by organometallic vapor phase epitaxy, *J. Cryst. Growth*, 124, 628 (1992).
15. J. E. Ayers, L. J. Schowalter, and S. K. Ghandhi, Post-growth thermal annealing of GaAs on Si (001) by organometallic vapor phase epitaxy, *J. Cryst. Growth*, 125, 329 (1992).
16. J.-F. He, H.-L. Wang, X.-J. Shang, M.-F. Li, Y. Zhu, L.-J. Wang, Y. Yu, H.-Q. Ni, Y.-Q. Xu, and Z.-C. Niu, GaAs-based long-wavelength InAs quantum dots on multi-step-graded InGaAs metamorphic buffer grown by molecular beam epitaxy, *J. Phys. D*, 44, 335102 (2011).
17. G. Salviati, C. Ferrari, L. Lazzarini, S. Franchi, A. Bosacchi, F. Taiariol, M. Mazzer, C. Zanotti-Fregonara, F. Romanato, and A. V. Drigo, TEM and x-ray diffraction studies of III-V lattice mismatched multilayers and superlattices, *Inst. Phys. Conf. Ser.*, 146, 337 (1995).
18. V. Krishnamoorthy, Y. W. Lin, and R. M. Park, Application of 'critical compositional difference' concept to the growth of low dislocation density ($<10^4/cm^2$) $In_xGa_{1-x}As$ ($x \leq 0.5$) on GaAs, *J. Appl. Phys.*, 72, 1752 (1992).
19. J. Knall, L. T. Romano, D. K. Biegelsen, R. D. Bringans, H. C. Chui, J. S. Harris Jr., D. W. Treat, and D. P. Bour, The use of graded InGaAs layers and patterned substrates to remove threading dislocations from GaAs on Si, *J. Appl. Phys.*, 76, 2697 (1994).
20. J.-I. Chyi, J.-L. Shieh, J.-W. Pan, and R.-M. Lin, Material properties of compositional graded $In_xGa_{1-x}As$ and $In_xAl_{1-x}As$ epilayers grown on GaAs substrates, *J. Appl. Phys.*, 79, 8367 (1996).
21. D. Gonzalez, D. Araujo, L. Gonzalez, Y. Gonzalez, G. Aragon, and R. Garcia, Advantages of thin interfaces in step-graded buffer structures, *Mater. Sci. Eng. B*, 44, 41 (1997).
22. X. Z. Shang, S. D. Wu, C. Liu, W. X. Wang, L. W. Guo, and Q. Huang, Low-temperature step-graded InAlAs/GaAs metamorphic buffer layers grown by molecular beam epitaxy, *J. Phys. D*, 39, 1800 (2006).
23. Y. Jeong, H. Choi, and T. Suzuki, Invalidity of graded buffers for InAs grown on GaAs (001)—A comparison between direct and graded-buffer growth, *J. Cryst. Growth*, 301–302, 235 (2007).
24. S.-I. Gozu, C. Hong, and S. Yamata, Low temperature high electron mobility in $In_{0.75}Ga_{0.25}As/In_{0.75}Al_{0.25}As$ modulation-doped heterostructures grown on GaAs substrate, *Jpn. J. Appl. Phys.*, Part 2, 37, L1501 (1998).
25. T. Mishima, K. Higuchi, M. Mori, and M. Kudo, High G_m $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ high electron mobility transistors grown lattice-mismatched on GaAs substrates, *J. Cryst. Growth*, 150, 1230 (1995).
26. L. J. Cui, Y. P. Zeng, B. Q. Wang, J. Wu, Z. P. Zhu, and L. Y. Lin, Rapid thermal annealing effects on step-graded InAlAs buffer layer and $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ metamorphic high electron mobility transistor structures on GaAs substrates, *J. Appl. Phys.*, 91, 2429 (2002).
27. S. Mendach, C. M. Hu, Ch. Heyn, S. Schnull, H. P. Oepen, R. Anton, and W. Hansen, Strain relaxation in high-mobility InAs inserted-channel heterostructures with metamorphic buffer, *Phys. E*, 13, 1204 (2002).
28. Ch. Heyn, S. Mendach, S. Lohr, S. Beyer, S. Schnull, and W. Hansen, Growth of shallow InAs HEMTs with metamorphic buffer, *J. Cryst. Growth*, 251, 832 (2003).
29. C. Lynch, R. Beresford, and E. Chason, Real-time stress evolution during growth of $In_xAl_{1-x}As/GaAs$ metamorphic buffer layers, *J. Vac. Sci. Technol. B*, 22, 1539 (2004).

30. F. Capotondi, G. Biasiol, D. Ercolani, V. Grillo, E. Carlino, F. Romanato, and L. Sorba, Strain induced effects on the transport properties of metamorphic InAlAs/InGaAs quantum wells, *Thin Solid Films*, 484, 400 (2005).
31. D. Lee, M. S. Park, Z. Tang, H. Luo, R. Beresford, and C. R. Wie, Characterization of metamorphic $In_xAl_{1-x}As$/GaAs buffer layers using reciprocal space mapping, *J. Appl. Phys.*, 101, 063523 (2007).
32. Z. Jiang, W. Wang, H. Gao, L. Liu, H. Chen, and J. Zhou, Strain relaxation and surface morphology of high indium content InAlAs metamorphic buffers with reverse step, *Appl. Surf. Sci.*, 254, 5241 (2008).
33. B. Landgraf, T. Slobodskyy, Ch. Heyn, and W. Hansen, Strain relaxation in metamorphic InAlAs buffers, *Mater. Sci. Eng. B*, 177, 762 (2012).
34. J.-H. Jang, G. Cueva, W. E. Hoke, P. J. Lemonias, P. Fay, and I. Adesida, Metamorphic graded bandgap InGaAs-InGaAlAs-InAlAs double heterojunction P-i-I-N photodiodes, *J. Lightwave Technol.*, 20, 507 (2002).
35. L. J. Mawst, J. D. Kirch, C.-C. Chang, T. Kim, T. Garrod, D. Botez, S. Ruder, et al., InGaAs/AlInAs strain-compensated superlattices grown on metamorphic buffer layers for low-strain, 3.6 μm-emitting quantum-cascade-laser active regions, *J. Cryst. Growth*, 370, 230 (2013).
36. M. Haupt, K. Kohler, P. Ganser, S. Emminger, S. Muller, and W. Rothemund, Growth of high quality $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures using strain relaxed $Al_xGa_yIn_{1-x-y}As$ buffer layers on GaAs, *Appl. Phys. Lett.*, 69, 412 (1996).
37. G. Balakrishnan, S. Huang, T. J. Rotter, A. Stintz, L. R. Dawson, K. J. Malloy, H. Xu, and D. L. Huffaker, 2.0 μm wavelength InAs quantum dashes grown on a GaAs substrate using a metamorphic buffer layer, *Appl. Phys. Lett.*, 84, 2058 (2004).
38. Y.-C. Xin, L. G. Vaughn, L. R. Dawson, A. Stintz, Y. Lin, L. F. Lester, and D. L. Huffaker, InAs quantum-dot GaAs-based lasers grown on AlGaAsSb metamorphic buffers, *J. Appl. Phys.*, 94, 2133 (2003).
39. J. Kirch, T. Garod, S. Kim, J. H. Park, J. C. Shin, L. J. Mawst, T. F. Kuech, X. Song, S. E. Babcock, I. Vurgaftman, J. R. Meyer, and T.-S. Kuan, $InAs_yP_{1-y}$ metamorphic buffer layers on InP substrates for mid-IR diode lasers, *J. Cryst. Growth*, 312, 1165 (2010).
40. S. S. Bui, H. P. Lee, and K. M. Yu, Metamorphic $InAs_yP_{1-y}$ (y = 0.30 − 0.75) and $Al_\delta In_{1-\delta}As_yP_{1-y}$ buffer layers on InP substrates, *Appl. Phys. Lett.*, 90, 212113 (2007).
41. J. A. Czaban, D. A. Thompson, and B. J. Robinson, Improved InAsP metamorphic layers grown on an InP substrate using underlying InP grown at low temperatures, *Semicond. Sci. Technol.*, 22, 408 (2007).
42. M. K. Hudait, M. Brenner, and S. A. Ringel, Metamorphic $In_{0.7}Al_{0.3}As/In_{0.69}Ga_{0.31}As$ thermovoltaic devices grown on graded $InAs_yP_{1-y}$ buffers by molecular beam epitaxy, *Solid State Electron.*, 53, 102 (2009).
43. E. Plis, P. Rotella, S. Raghavan, L. R. Dawson, S. Krishna, D. Le, and C. P. Morath, Growth of room-temperature arsenic free infrared photovoltaic detectors on GaSb substrate using metamorphic InAlSb digital alloy buffer layers, *Appl. Phys. Lett.*, 82, 1658 (2003).
44. M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, Dislocation morphology in graded heterojunctions: $GaAs_{1-x}P_x$, *J. Mater. Sci.*, 4, 223 (1969).
45. J. Kirch, T. W. Kim, J. Konen, L. J. Mawst, T. F. Kuech, and T.-S. Kuan, Effects of antimony (Sb) incorporation on MOVPE grown $InAs_yP_{1-y}$ metamorphic buffer layers on InP substrates, *J. Cryst. Growth*, 315, 96 (2011).
46. M. A. G. Halliwell, M. H. Lyons, and M. J. Hill, The interpretation of x-ray rocking curves from III-V semiconductor device structures, *J. Cryst. Growth*, 68, 523 (1984).
47. W. J. Bartels, J. Hornstra, and D. J. W. Lobeek, X-ray diffraction of multilayers and superlattices, *Acta Cryst. A*, 42, 539 (1986).
48. C. R. Wie, T. A. Tombrello, and T. Vreeland Jr., Dynamical x-ray diffraction from nonuniform crystalline films: Application to x-ray rocking curve analysis, *J. Appl. Phys.*, 59, 3743 (1986).
49. P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Dynamical x-ray diffraction from ZnS_ySe_{1-y}/GaAs (001) multilayers and superlattices with dislocations, *J. Electron. Mater.*, 41, 2846 (2012).

50. P. B. Rago and J. E. Ayers, Dynamical x-ray diffraction from $In_xGa_{1-x}As$ heterostructures with dislocations, *J. Electron. Mater.*, 42, 2450 (2012).
51. T. L. Song, S. J. Chua, E. A. Fitzgerald, P. Chen, and S. Tripathy, Strain relaxation in graded InGaN/GaN epilayers grown on sapphire, *Appl. Phys. Lett.*, 83, 1545 (2003).
52. J. Chen, J. M. Fernandez, J. C. P. Chang, K. L. Kavanagh, and H. H. Wieder, Modulation-doped $In_{0.3}Ga_{0.7}As/In_{0.29}Al_{0.71}As$ heterostructures grown on GaAs by step grading, *Semicond. Sci. Technol.*, 7, 601 (1992).
53. S. Gozu, K. Tsuboki, M. Hayashi, C. Hong, and S. Yamada, Very high electron mobilities at low temperatures in $In_xGa_{1-x}As/In_yAl_{1-y}As$ HEMTs grown lattice-mismatched on GaAs substrates, *J. Cryst. Growth*, 201–202, 749 (1999).
54. D. E. Grider, S. E. Swirhun, D. H. Narum, A. I. Akinwande, T. E. Nohava, W. R. Stuart, and P. Joslyn, Metamorphic $In_xGa_{1-x}As/In_xAl_{1-x}As$ heterostructure field effect transistors grown on GaAs (001) substrates using molecular beam epitaxy, *J. Vac. Sci. Technol. B*, 8, 301 (1990).
55. Y. Huo, H. Lin, R. Chen, M. Makarova, Y. Rong, M. Li, T. I. Kamins, J. Vuckovic, and J. S. Harris, Strong enhancement of direct transition photoluminescence with highly tensile-strained Ge grown by molecular beam epitaxy, *Appl. Phys. Lett.*, 98, 011111 (2011).
56. Z. Mi, C. Wu, J. Yang, and P. Bhattacharya, Molecular beam epitaxial growth and characteristics of 1.52 μm metamorphic InAs quantum dot lasers on GaAs, *J. Vac. Sci. Technol. B*, 26, 1153 (2008).
57. J.-F. He, H.-L. Wang, X.-J. Shang, M.-F. Li, Y. Zhu, L.-J. Wang, Y. Yu, H.-Q. Ni, Y.-Q. Xu, and Z.-C. Niu, GaAs-based long-wavelength InAs quantum dots on multi-step-graded InGaAs metamorphic buffer grown by molecular beam epitaxy, *J. Phys. D*, 44, 335102 (2011).
58. L. Shen, H. H. Wieder, and W. S. C. Chang, Low temperature grown thin InAlAs step graded buffers for application on optical modulator at 1.3 μm, *Mater. Res. Soc. Symp.*, 379, 297 (1997).
59. S.-G. Ihn, S.-J. Jo, K.-H. Oh, T. W. Kim, and J.-I. Song, Crystalline quality improvement of $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ heterostructure on InAlAs/GaAs metamorphic buffer by post-growth rapid thermal annealing, in *2005 International Conference on Indium Phosphide and Related Materials*, 216 (2005).
60. J. E. Ayers, S. K. Ghandhi, and L. J. Schowalter, Crystallographic tilting of heteroepitaxial layers, *J. Cryst. Growth*, 113, 430 (1991).
61. C. J. Nuese, J. J. Tietjen, J. J. Gannon, and H. F. Gossenberger, Electroluminescence of vapor-grown GaAs and $GaAs_{1-x}P_x$ diodes, *Trans. AIME*, 242, 400 (1968).
62. I. Tangring, H. Q. Ni, B. P. Wu, D. H. Wu, Y. H. Xiong, S. S. Huang, Z. C. Niu, S. M. Wang, Z. H. Lai, and A. Larsson, 1.58 μm InGaAs quantum well laser on GaAs, *Appl. Phys. Lett.*, 91, 221101 (2007).
63. Y. Song, S. Wang, I. Tangring, Z. Lai, and M. Sadeghi, Effects of doping and grading slope on surface and structure of metamorphic InGaAs buffers on GaAs substrates, *J. Appl. Phys.*, 106, 123531 (2009).
64. S. I. Molina, F. J. Pacheco, D. Araujo, R. Garcia, A. Sacedon, E. Calleja, Z. Yang, and P. Kidd, Strain relief in linearly graded composition buffer layers: A design scheme to grow dislocation-free (<102 7 cm−2) and unstrained epilayers, *Appl. Phys. Lett.*, 65, 2460 (1994).
65. A. Sacedon, F. Gonzalez-Sanz, E. Calleja, E. Munoz, S. I. Molina, F. J. Pacheco, D. Araujo, et al., Design of InGaAs linear graded buffer structures, *Appl. Phys. Lett.*, 66, 3334 (1995).
66. S.-D. Kim, S. M. Lord, and J. S. Harris Jr., Strain relaxation in compositionally graded epitaxial layers, *J. Vac. Sci. Technol. B*, 14, 642 (1996).
67. M. T. Bulsara, V. Yang, A. Thilderkvist, E. A. Fitzgerald, K. Hausler, and K. Eberl, Graded $In_xGa_{1-x}As$/GaAs 1.3 μm wavelength light emitting diode structures grown with molecular beam epitaxy, *J. Appl. Phys.*, 83, 592 (1998).
68. I. Tangring, S. Wang, M. Sadeghi, and A. Larsson, 1.27 μm metamorphic InGaAs quantum well lasers on GaAs substrates, *Electron. Lett.*, 42, 691 (2006).
69. I. Tangring, S. M. Wang, X. R. Zhu, A. Larsson, Z. H. Lai, and M. Sadeghi, Manipulation of strain relaxation in metamorphic heterostructures, *Appl. Phys. Lett.*, 90, 071904 (2007).
70. K. E. Lee and E. A. Fitzgerald, High-quality metamorphic compositionally graded InGaAs buffers, *J. Cryst. Growth*, 312, 250 (2010).

71. L. Yang, M. T. Bulsara, K. E. Lee, and E. A. Fitzgerald, Compositionally-graded InGaAs-InGaP alloys and GaAsSb alloys for metamorphic InP on GaAs, *J. Cryst. Growth*, 324, 103 (2011).
72. J. A. Olsen, E. L. Hu, S. R. Lee, I. J. Fritz, A. J. Howard, B. E. Hammons, and J. Y. Tsao, X-ray reciprocal-space mapping of strain relaxation and tilting in linearly graded InAlAs buffers, *J. Appl. Phys.*, 79, 3578 (1996).
73. M. Dammann, M. Chertouk, W. Jantz, K. Kohler, and G. Weimann, Reliability of InAlAs/InGaAs HEMTs grown on GaAs substrate with metamorphic buffer, *Microelectron. Reliab.*, 40, 1709 (2000).
74. E. S. Semenova, A. E. Zhukov, S. S. Mikhrin, A. Y. Egorov, V. A. Odnoblyudov, A. P. Vasil'ev, E. V. Nikitina, et al., Metamorphic growth for application in long-wavelength (1.3–1.55 μm) lasers and MODFET-type structures on GaAs substrates, *Nanotechnology*, 15, S283 (2004).
75. A. Cavus, C. Monier, C. Cox, D. Pascua, A. Gutierrez-Aitken, A. Noori, S. Hayashi, and M. Goorsky, Development of 6.00 Å graded metamorphic buffer layers and high performance $In_{0.86}Al_{0.14}As/In_{0.86}Ga_{0.14}As$ heterojunction bipolar transistor devices, *J. Vac. Sci. Technol. B*, 24, 1492 (2006).
76. Y. Zhang, Y. Gu, Z. Tian, A. Li, X. Zhu, and Y. Zheng, Wavelength extended 2.4 μm heterojunction InGaAs photodiodes with InAlAs cap and linearly graded buffer layers suitable for both front and back illuminations, *Infrared Phys. Technol.*, 51, 316 (2008).
77. H. Choi, Y. Jeong, J. Cho, and M. H. Jeon, Effectiveness of non-linear graded buffers for In(Ga,Al)As metamorphic layers grown on GaAs (001), *J. Cryst. Growth*, 311, 1091 (2009).
78. G. B. Galiev, I. S. Vasil'evskii, S. S. Pushkarev, E. A. Klimov, R. M. Imamov, P. A. Buffat, B. Dwir, and E. I. Suvorova, Metamorphic InAlAs/InGaAs/InAlAs/GaAs HEMT heterostructures containing strained superlattices and inverse steps in the metamorphic buffer, *J. Cryst. Growth*, 366, 55 (2013).
79. W. E. Hoke, T. D. Kennedy, A. Torabi, C. S. Whelan, P. F. Marsh, R. E. Leoni, S. M. Lardizabal, et al., Properties of metamorphic materials and device structures on GaAs substrates, *J. Cryst. Growth*, 251, 804 (2003).
80. L. J. Mawst, J. D. Kirch, C.-C. Chang, T. Kim, T. Garrod, D. Botez, S. Ruder, et al., InGaAs/AlInAs strain-compensated superlattices grown on metamorphic buffer layers for low-strain, 3.6 μm-emitting quantum-cascade-laser active regions, *J. Cryst. Growth*, 370, 230 (2013).
81. A. Wakita, H. Rohden, V. Robbins, N. Mll, Chung-Yi Su, A. Nagy, and D. Basile, Low-noise bias reliability of AlInAs/GaInAs modulation-doped field effect transistors with linearly graded low-temperature buffer layers grown on GaAs substrates, *Jpn. J. Appl. Phys.*, 38, 1186 (1999).
82. S. J. Jo, S.-G. Ihn, G.-J. Lee, D.-H. Lee, and J.-I. Song, Carrier lifetime of low-temperature grown $In_{0.53}Ga_{0.47}As$ on GaAs using an InGaAlAs metamorphic buffer, in *2005 International Conference on Indium Phosphide and Related Materials*, 175 (2005).
83. J. L. Liu, K. L. Wang, C. D. Moore, M. S. Goorsky, T. Borca-Tasciuc, and G. Chen, Experimental study of a surfactant-assisted SiGe graded layer and a symmetrically strained Si/Ge superlattice for thermoelectric applications, *Thin Solid Films*, 369, 121 (2000).
84. K. L. Lew, S. F. Yoon, H. Tanato, K. P. Chen, C. L. Dohrman, D. M. Isaacson, and E. A. Fitzgerald, InGaP/GaAs heterojunction bipolar transistor grown on Si substrate with SiGe graded buffer layer, *Electron. Lett.*, 44, 243 (2008).
85. J. Endres, S. Danis, and G. Bauer, The misfit dislocation density profile in graded SiGe/Si(001) layers prepared at different temperatures, *J. Phys. Cond. Matt.*, 25, 175802 (2013).
86. H. Yang, H. Wang, K. Radhakrishnan, and C. L. Tan, Thermal resistance of metamorphic InP-based HBTs on GaAs substrates using a linearly graded $In_xGa_{1-x}P$ metamorphic buffer, *IEEE Trans. Electron Dev.*, 51, 1221 (2004).
87. J.-H. Tsai, Study of metamorphic co-integrated heterostructure bipolar and field-effect transistors (BiFETs), *J. Electrochem. Soc.*, 158, H889 (2011).
88. A. Zakaria, R. R. King, M. Jackson, and M. S. Goorsky, Comparison of arsenide and phosphide based graded buffer layers used in inverted metamorphic solar cells, *J. Appl. Phys.*, 112, 024907 (2012).

89. V. A. Odnoblyudov and C. W. Tu, Room-temperature yellow-amber emission from InGaP quantum wells grown on an InGaP metamorphic buffer layer on GaP (100) substrates, *J. Cryst. Growth*, 279, 25 (2005).
90. J. F. Ocampo, B. Bertoli, P. B. Rago, E. N. Suarez, D. Shah, F. C. Jain, and J. E. Ayers, Asymmetric dislocation densities in forward-graded ZnS_ySe_{1-y}/GaAs (001) heterostructures, *J. Electron. Mater.*, 39, 391 (2010).
91. T. Hosoda, D. Wang, G. Kipshidze, W. L. Sarney, and L. Shterengas, 3 μm diode lasers grown on (Al)GaInSb compositionally graded metamorphic buffer layers, *Semicond. Sci. Technol.*, 27, 055011 (2012).
92. M. Haupt, K. Kohler, P. Ganser, S. Muller, and W. Rothemund, Molecular beam epitaxy of $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures on metamorphic $Al_xGa_yIn_{1-x-y}As$ buffer layers, *J. Cryst. Growth*, 175–176, 1028 (1997).
93. B. Bertoli, E. N. Suarez, F. C. Jain, and J. E. Ayers, Dislocations and strain relief in reverse-graded semiconductor layers, *Semicond. Sci. Technol.*, 24, 125006 (2009).
94. L. H. Wong, J. P. Liu, C. C. Wong, C. Ferraris, T. J. White, L. Chan, D. K. Sohn, and L. C. Hsia, Thermal stability of a reverse-graded SiGe buffer layer for growth of relaxed SiGe epitaxy, *Electrochem. Solid State Lett.*, 9, G114 (2006).
95. E. A. Fitzgerald, Y.-H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, A. R. Kortan, F. A. Thiel, and B. E. Weir, Relaxed Ge_xSi_{1-x} structures for III-V integration with Si and high mobility two-dimensional electron gases in Si, *J. Vac. Sci. Technol. B*, 10, 1807 (1992).
96. J. E. Ayers, Critical layer thickness in compositionally-graded semiconductor layers with non-zero interfacial mismatch, *Semicond. Sci. Technol.*, 23, 045018 (2008).
97. E. A. Fitzgerald, Dislocations in strained-layer epitaxy: Theory, experiment, and applications, *Mater. Sci. Rep.*, 7, 87 (1991).
98. J. E. Ayers, Multiple critical layer thicknesses in retro-graded heterostructures, *Appl. Phys. Lett.*, 92, 102104 (2008).
99. G. Belenky, D. Wang, Y. Lin, D. Donetsky, G. Kipshidze, L. Shterengas, D. Westerfeld, W. L. Sarney, and S. P. Svensson, Metamorphic InAsSb/AlInAsSb heterostructures for opto-electronic applications, *Appl. Phys. Lett.*, 102, 111108 (2013).
100. B. Lee, J. H. Baek, J. H. Lee, S. W. Choi, S. D. Jung, W. S. Han, and E. H. Lee, Optical properties of InGaAs linear graded buffer layers on GaAs grown by metalorganic chemical vapor deposition, *Appl. Phys. Lett.*, 68, 2973 (1996).
101. J. F. Ocampo, E. Suarez, F. C. Jain, and J. E. Ayers, Overshoot graded layers for mismatched heteroepitaxial devices, *J. Electron. Mater.*, 37, 1035 (2008).
102. Y. Cordier, J.-M. Chauveau, D. Ferre, and J. Dipersio, Comparison of $In_{0.33}Ga_{0.67}As/In_{0.34}Al_{0.66}As$ on GaAs metamorphic high electron mobility transistors grown by molecular beam epitaxy with normal and inverse step on linear graded buffer layers, *J. Vac. Sci. Technol. B*, 18, 2513 (2000).
103. T. Kujofsa and J. E. Ayers, Equilibrium lattice relaxation and misfit dislocation in continuously- and step-graded $In_xGa_{1-x}As$/GaAs (001) and $GaAs_{1-y}P_y$/GaAs (001) metamorphic buffer layers, *Intl. J. High Speed Electron. Syst.*, (24), 1520009 (2015).
104. J. J. Tietjen, J. I. Pankove, I. J. Hegyi, and H. Nelson, Vapor-phase growth of $GaAs_{1-x}P_x$, room-temperature injection lasers, *Trans. AIME*, 239, 385 (1967).
105. C. J. Nuese, J. J. Tietjen, J. J. Gannon, and H. F. Gossenberger, Electroluminescence of vapor-grown GaAs and $Gas_{1-x}P_x$ diodes, *Trans. AIME*, 242, 400 (1968).
106. D. Richman and J. J. Tietjen, Rapid vapor phase growth of high-resistivity GaP for electro-optic modulators, *Trans. AIME*, 239, 418 (1967).
107. R. H. Saul, Effect of a $GaAs_xP_{1-x}$ transition zone on the perfection of GaP crystals grown by deposition onto GaAs substrates, *J. Appl. Phys.*, 40, 3273 (1969).
108. E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, Dislocation dynamics in relaxed graded composition semiconductors, *Mater. Sci. Eng. B*, 67, 53 (1999).
109. F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, Mechanism and conditions for anomalous strain relaxation in graded thin films and superlattices, *J. Appl. Phys.*, 71, 4230 (1992).

110. F. K. LeGoues, P. M. Mooney, and J. O. Chu, Crystallographic tilting resulting from nucleation limited relaxation, *Appl. Phys. Lett.*, 62, 140 (1993).
111. F. K. LeGoues, P. M. Mooney, and J. Tersoff, Measurement of the activation barrier to nucleation of dislocations in thin films, *Phys. Rev. Lett.*, 71, 396 (1993); *Phys. Rev. Lett.*, 71, 3234 (1993).
112. H. Nagai, Structure of vapor-deposited $Ga_xIn_{1-x}As$ crystals, *J. Appl. Phys.*, 45, 3789 (1974).
113. K. Muraki, S. Fukatsu, Y. Shiraki, and R. Ito, Surface segregation of In atoms during molecular beam epitaxy and its influence on the energy levels in InGaAs/GaAs quantum wells, *Appl. Phys. Lett.*, 61, 557 (1992).
114. A. S. Wakita, H. Rohdin, V. M. Robbins, N. Moll, C.-Y. Su, A. Nagy, and D. P. Basile, Low-noise bias reliability of AlInAs/GaInAs MODFETs with linearly graded low-temperature buffer layers grown on GaAs substrates, in *10th International Conference on Indium Phosphide and Related Materials*, TuP-38 (1998).
115. H. Y. Liu, Y. Qiu, C. Y. Jin, T. Walther, and A. G. Cullis, 1.55 μm InAs quantum dots grown on a GaAs substrate using a GaAsSb metamorphic buffer layer, *Appl. Phys. Lett.*, 92, 111906 (2008).
116. B. Bertoli, D. Sidoti, S. Xhurxhi, T. Kujofsa, S. Cheruku, J. P. Correa, P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Equilibrium strain and dislocation density in exponentially graded $Si_{1-x}Ge_x$/Si (001), *J. Appl. Phys.*, 108, 113525 (2010).
117. D. Sidoti, S. Xhurxhi, T. Kujofsa, S. Cheruku, J. Reed, B. Bertoli, P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Critical layer thickness in exponentially graded heteroepitaxial layers, *J. Electron. Mater.*, 29, 1140 (2010).
118. I. Mathews, D. O'Mahony, A. Gocalinska, M. Manganaro, E. Palucchi, M. Schmidt, A. P. Morrison, and B. Corbett, InAlAs solar cell on a GaAs substrate employing a graded $In_xGa_{1-x}As$-InP metamorphic buffer layer, *Appl. Phys. Lett.*, 102, 033906 (2013).
119. P. M. J. Maree, J. C. Barbour, J. F. van der Veen, K. L. Kavanough, C. W. T. Bulle-Lieuwma, and M. P. A. Viegers, Generation of misfit dislocations in semiconductors, *J. Appl. Phys.*, 62, 4413 (1987).
120. A. V. Drigo, A. Aydinli, A. Carnera, F. Genova, C. Rigo, C. Ferrari, P. Franzosi, and G. Salviati, On the mechanisms of strain release in molecular-beam-epitaxy-grown $In_xGa_{1-x}As$/GaAs single heterostructures, *J. Appl. Phys.*, 66, 1975 (1989).
121. C. Ferrari, P. Franzosi, L. Lazzarini, G. Salviati, M. Berti, A. V. Drigo, M. Mazzer, F. Romanato, M. R. Bruni, and M. G. Simeone, Mechanisms of strain release in molecular beam epitaxy grown InGaAs/GaAs buffer heterostructures, *Mater. Sci. Eng. B*, 28, 510 (1994).
122. T. Kujofsa and J. E. Ayers, Lattice relaxation and misfit dislocations in nonlinearly graded In_xGa_1–xAs/GaAs (001) and $GaAs_1$–yP_y/GaAs (001) metamorphic buffer layers, *J. Vac. Sci. Technol. B*, 32, 031205 (2014).
123. T. Kujofsa and J. E. Ayers, Design of nonlinear metamorphic buffer layers for lattice-mismatched $In_xGa_{1-x}As$/GaAs (001) semiconductor devices, *J. Vac. Sci. Technol. B*, 33, 052206 (2015).
124. S. Xhurxhi, F. Obst, D. Sidoti, B. Bertoli, T. Kujofsa, S. Cheruku, J. P. Correa, P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, S-graded buffer layers for lattice-mismatched heteroepitaxial devices, *J. Electron. Mater.*, 40, 2348 (2011).
125. T. Kujofsa, A. Antony, S. Xhurxhi, F. Obst, D. Sidoti, B. Bertoli, S. Cheruku, et al., Design of S-graded buffer layers for metamorphic ZnS_ySe_{1-y}/GaAs (001) semiconductor devices, *J. Electron. Mater.*, 42, 3408 (2013).
126. T. Kujofsa and J. E. Ayers, Threading dislocations in S-graded ZnS_ySe_{1-y}/GaAs (001) metamorphic buffer layers, *Intl. J. High Speed Electron. Syst.*, 23, 1420005 (2014).
127. R. Fischer, H. Morkoc, D. A. Nueman, H. Zabel, C. Choi, N. Otsuka, M. Longerbone, and L. P. Erickson, Material properties of high-quality GaAs epitaxial layers grown on Si substrates, *J. Appl. Phys.*, 60, 1640 (1986).
128. M. Yamaguchi, T. Nishioka, and M. Sugo, Analysis of strained-layer superlattice effects on dislocation density reduction in GaAs on Si substrates, *Appl. Phys. Lett.*, 54, 24 (1989).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

7

Characterization of Heteroepitaxial Layers

7.1 Introduction

Numerous and diverse characterization techniques have been used for the evaluation of heteroepitaxial semiconductors, and have enabled the advancement of the field to its present state. It would be impossible to describe all of them here. Instead, this chapter emphasizes some of the most commonly used techniques, such as x-ray diffraction (XRD), electron diffraction, electron microscopy, crystallographic etching, and photoluminescence (PL). These basic methods have contributed greatly to our current understanding of heteroepitaxy. Some have also been adapted as routine characterization methods for commercial production of heteroepitaxial materials and devices.

High-resolution x-ray diffraction (HRXRD) is the most widely used technique for the *ex situ* characterization of heteroepitaxial layers. HRXRD is nondestructive and yields a wealth of structural information, including the lattice constants and strains, crystallographic orientation, and defect densities. Reciprocal space mapping and rocking curve measurements at different azimuths can also be used to study the asymmetries in the defect densities on different slip systems. In multilayer device structures, dynamical rocking curve simulations can be used to extract the thicknesses, lattice constants, and compositions in the individual layers. The versatility and nondestructive nature of HRXRD have led to its common use in production environments, as well as in basic studies.

Electron diffraction methods have also been used to a great extent, for the study of surface structures, surface reconstructions, nucleation, and growth modes. The commonly used methods include reflection high-energy electron diffraction (RHEED) and low-energy electron diffraction (LEED), and their variants. The high-vacuum environment used for molecular beam epitaxy (MBE) growth allows the *in situ* use of electron diffraction techniques, which provide valuable feedback during the growth process.

Microscopic methods include optical microscopy (OM), scanning electron microscopy (SEM), transmission electron microscopy (TEM), atomic force microscopy (AFM), and scanning tunneling electron microscopy (STEM). Most of these are used to image the surface, and thus characterize the surface morphology and growth mode. They may also be used in conjunction with crystallographic etches to determine defect densities. AFM has sufficient resolution for the study of the surface atomic structure. TEM enables the imaging of dislocations and other crystal defects within the volume of the crystal; it is therefore important in the study of dislocations and lattice relaxation in mismatched heteroepitaxial layers.

Crystallographic etching techniques are used routinely to determine threading dislocation densities in single-crystal substrates as well as heteroepitaxial layers. Crystallographic etches, due to their surface-sensitive etch rates, reveal pits at the points of emergence for

defects. Subsequent microscopic inspection can be used to determine the dislocation density. In some materials, different types of dislocation defects can be distinguished by their characteristic pit shapes. Moreover, crystal orientation may be determined by the alignment of oval pits.

PL is commonly employed *ex situ* to assess the suitability of heteroepitaxial structures for optoelectronic devices such as light-emitting diodes (LEDs) and laser diodes. A wealth of information may be obtained from PL spectra, especially by taking measurements at different temperatures or with different excitation wavelengths or intensities. Much of this information is particularly useful in studies of doping, which are beyond the scope of this book and will not be elaborated here. It is also possible to use PL for the determination of structural information, such as compositions and strains in heteroepitaxial layers. This is not commonly done, though, because the analysis involves many nonstructural factors and is far more complex than structural characterization by HRXRD. An exception is the case of quaternary layers, for which PL is commonly used in conjunction with HRXRD for determination of the composition and strain. Relative intensities measured by PL are useful in assessing the influence of crystal defects on the minority carrier lifetime. An important extension of this is photoluminescence microscopy (PLM), which can be used to image individual dislocations for the study of lattice relaxation in heteroepitaxial semiconductors. Related luminescence imaging techniques can be used similarly; these include cathodoluminescence (CL) and electroluminescence.

This chapter describes the principles of these commonly used characterization techniques and their application to heteroepitaxial layers.

7.2 X-Ray Diffraction

High-resolution x-ray diffraction (HRXRD) is important in the structural characterization of heteroepitaxial layers, revealing lattice constants, strains, crystallographic orientation, and defect densities. It can also be used in conjunction with dynamic simulations for the extraction of the compositions, strains, and thicknesses of individual layers in multilayer device structures.

A typical high-resolution x-ray diffractometer is illustrated in Figure 7.1. The source of x-rays is usually an x-ray tube, producing a divergent beam with a broad spectrum. The beam is conditioned (limited in both angular divergence and wavelength spread) by four diffracting surfaces arranged in a *Bartels monochromator*. The conditioned beam is then diffracted by the specimen crystal and measured using a scintillation detector. In a rocking

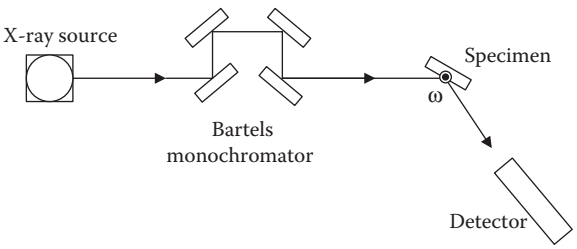


FIGURE 7.1
High-resolution x-ray diffractometer.

curve measurement, the specimen is rotated about the ω axis (which is perpendicular to the plane of the page and passes through the point where the beam strikes the sample). The diffracted intensity is measured as a function of the specimen angle ω. The positions, intensities, and widths of the intensity peaks in this diffraction profile (or "rocking curve") are used to characterize the structural properties of the specimen crystal. Application of this method therefore requires an understanding of how the rocking curve relates to the specimen crystal structure. This section describes the basic principles needed for the application of this technique to the characterization of heteroepitaxial layers, as well as some practical aspects of diffractometer instruments.

7.2.1 Positions of Diffracted Beams

Diffraction from a three-dimensional crystal is the constructive interference of waves scattered by the atoms in the lattice. A necessary condition for diffraction is that the path length difference for beams scattered from different atoms be an integral multiple of the x-ray wavelength. This condition may be stated in two equivalent ways: the Bragg equation, which is a geometric equation in real space, and the Laue equations, which are the equivalent condition expressed in reciprocal space. The positions of the atoms within the unit cell impose further conditions on diffraction.

7.2.1.1 Bragg Equation

The Bragg equation for diffraction may be understood with the aid of Figure 7.2. Here, an x-ray beam is incident on a set of crystal planes with separation d. If the angles of incidence and reflection are equal to θ (specular reflection), then the path difference Δ between the beams a and b is $2d \sin\theta$. The condition for constructive interference is $\Delta = n\lambda$, where n is an integer and λ is the x-ray wavelength. Thus, the condition for diffraction is

$$2d \sin\theta_B = n\lambda, \tag{7.1}$$

where n is the order of the reflection. This is the Bragg equation,[1] and θ_B is the Bragg angle.

7.2.1.1.1 Bragg Equation in Cubic Semiconductors

For a cubic crystal with lattice constant a, the spacing of the (hkl) planes is

$$d(hkl) = a\left(h^2 + k^2 + l^2\right)^{-1/2}, \tag{7.2}$$

and the hkl Bragg angle is

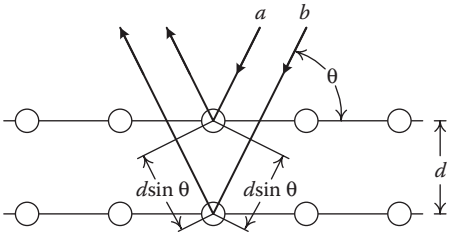


FIGURE 7.2
Bragg condition for diffraction.

$$\theta_B(hkl) = \sin^{-1}\left[\lambda\left(h^2+k^2+l^2\right)^{1/2}/(2a)\right]. \tag{7.3}$$

For a substrate with a (*uvw*) surface orientation, the angle between the (*hkl*) planes and the (*uvw*) surface is given by

$$\phi(hkl) = \cos^{-1}\left(\frac{uh+vk+wl}{\sqrt{\left(h^2+k^2+l^2\right)\left(u^2+v^2+w^2\right)}}\right). \tag{7.4}$$

7.2.1.1.2 Bragg Equation in Hexagonal Semiconductors[2]

For a hexagonal crystal with lattice constants *a* and *c*, the spacing of the (*hkil*) planes (also sometimes referred to as the (*hkl*) planes, because *i* is not an independent index, but is given by $i = -(h + k)$) is

$$d(hkil) = \left(\frac{h^2+hk+k^2}{3a^2/4}+\frac{l^2}{c^2}\right)^{-1/2}. \tag{7.5}$$

The *hkil* Bragg angle is

$$\theta_B(hkil) = \sin^{-1}\left[\frac{\lambda}{2}\left(\frac{h^2+hk+k^2}{3a^2/4}+\frac{l^2}{c^2}\right)^{1/2}\right]. \tag{7.6}$$

The angle between the *hkil* planes and the *uvzw* surface plane is

$$\phi(h,k,i,l) = \cos^{-1}\left[d(hkl)\cdot d(uvw)\cdot\left(\left[uh+vk+\frac{1}{2}(vh+uk)\right]\cdot\frac{4}{3a^2}+\frac{wl}{c^2}\right)\right]. \tag{7.7}$$

7.2.1.2 Reciprocal Lattice and the von Laue Formulation for Diffraction

For a crystal lattice in real space, there is a corresponding lattice in reciprocal space. This reciprocal lattice finds application in the analysis of wave interactions with the crystal, as in diffraction. The reciprocal lattice points lie at the tips of all wave vectors K that yield plane waves with the periodicity of the real lattice.[3] Thus, the reciprocal lattice vectors **K** are the set of vectors satisfying

$$\exp\{i(\mathbf{K}\cdot\mathbf{R})\} = 1 \tag{7.8}$$

for all real lattice vectors **R**, which are generated by

$$\mathbf{R} = m\mathbf{a} + n\mathbf{b} + o\mathbf{c}, \tag{7.9}$$

where:

a, **b**, and **c** are the primitive translation vectors of the real lattice
m, *n*, and *o* are integers

The reciprocal lattice vectors are generated by

$$\mathbf{K} = h\mathbf{A} + k\mathbf{B} + l\mathbf{C}, \tag{7.10}$$

where h, k, and l are integers. The primitive translation vectors of the reciprocal lattice are given by

$$\mathbf{A} = 2\pi(\mathbf{b}\times\mathbf{c})/(\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}), \tag{7.11}$$

$$\mathbf{B} = 2\pi(\mathbf{c}\times\mathbf{a})/(\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}), \tag{7.12}$$

and

$$\mathbf{C} = 2\pi(\mathbf{a}\times\mathbf{b})/(\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}). \tag{7.13}$$

Any arbitrary set of primitive vectors for the crystal lattice will reproduce a unique reciprocal lattice. Whereas the crystal lattice vectors have units of length in real space, the reciprocal lattice vectors have units of reciprocal length in the associated reciprocal space. The reciprocal of the face-centered cubic lattice with lattice constant a is body-centered cubic with cube side $4\pi/a$.

The von Laue condition for diffraction may be understood using Figure 7.3. Incident radiation with wave vector k is scattered at two points of the lattice that are displaced by a vector d. For constructive interference of the scattered waves, the path difference must be an integral number of wavelengths, or

$$\mathbf{d}\cdot(\mathbf{n}-\mathbf{n}') = m\lambda, \tag{7.14}$$

where:

$\mathbf{n}$ and $\mathbf{n}'$ are the unit vectors parallel to $\mathbf{k}$ and $\mathbf{k}'$, respectively
m is an integer

Multiplying both sides by $2\pi/\lambda$, which is the magnitude of the wave vectors, we obtain

$$\mathbf{d}\cdot(\mathbf{k}-\mathbf{k}') = (2\pi)m. \tag{7.15}$$

The condition for diffraction from the crystal is that Equation 7.13 hold for all pairs of scatterers, or

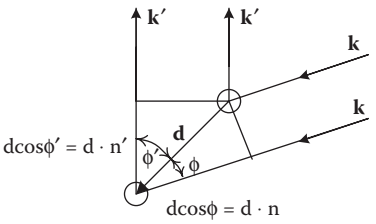


FIGURE 7.3
von Laue condition for diffraction.

$$\mathbf{R}\cdot(\mathbf{k}-\mathbf{k}')=(2\pi)m, \tag{7.16}$$

for all lattice vectors R. This may be written in the equivalent form

$$\exp\{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}\}=1. \tag{7.17}$$

Combining this equation with the defining equations for the reciprocal lattice, we arrive at the von Laue condition for diffraction. That is, constructive interference will occur if the scattering vector $\Delta\mathbf{k} = (\mathbf{k}–\mathbf{k}')$ is equal to a reciprocal lattice vector, or

$$(\mathbf{k}-\mathbf{k}')=h\mathbf{A}+k\mathbf{B}+l\mathbf{C}(\mathbf{k}-\mathbf{k}')=h\mathbf{A}+k\mathbf{B}+l\mathbf{C}, \tag{7.18}$$

where the integers h, k, and l are the indices for the reflection.

7.2.1.3 Ewald Sphere

The diffraction condition in reciprocal space may be represented by the geometric construction of Ewald, as shown in Figure 7.4. The vector $\boldsymbol{k}$ is drawn in the direction of the incident x-ray beam and terminating on the origin. A sphere of radius $k = |\boldsymbol{k}|$ is drawn with its center at the other end of the vector $\boldsymbol{k}$ and passing through the origin. This is the Ewald sphere. A reflection will be excited if any reciprocal lattice point G lies on the surface of the sphere.

7.2.2 Intensities of Diffracted Beams

Whereas the positions of the diffracted beams depend only on the unit cell dimensions, the intensities of Bragg reflections depend on how the x-rays are scattered within each unit cell, by the electrons surrounding individual atoms.

In this section, the relative intensities of the Bragg reflections are determined, starting with a description of the scattering of x-rays by a single isolated electron. The individual atoms are treated as ensembles of scattering electrons. Then it is shown that the diffracted intensity is determined by the interference of scattered x-rays from all of the atoms in the unit cell.

These considerations lead to many important features of x-ray diffraction (XRD), which pertain to its use as a characterization tool. These include the different intensities observed

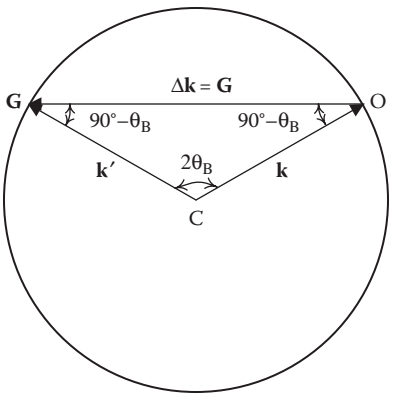


FIGURE 7.4
Ewald sphere construction.

for different reflections from a single crystal, the finite angular widths of diffraction profiles from perfect crystals, and the forbidden reflections.

7.2.2.1 Scattering of X-Rays by a Single Electron

X-rays are scattered in all directions by a single electron, with the scattered intensity strongly dependent on the scattering angle α. This dependence was derived by J. J. Thompson and is

$$I = I_0\left(\frac{\mu_0}{4\pi}\right)^2\left(\frac{q^4}{m^2r^2}\right)\sin^2\alpha, \tag{7.19}$$

where:

- I is the scattered intensity at a distance r and an angle α, $\mu_0 = 4\pi \times 10^{-7}$ Hm^{-1}
- q is the electronic charge, 1.602×10^{-19} C
- m is the electron rest mass, 9.11×10^{-31} kg
- α is the angle between the scattering direction and the direction of acceleration for the electron, and therefore depends on the polarization of the x-ray beam

X-ray beams obtained from x-ray tubes are unpolarized, so that the electric vector **E** has a random orientation in the plane perpendicular to the beam. Referring to Figure 7.5, an unpolarized x-ray beam issuing from point N encounters an electron at point O and the scattered beam is observed at point P. **E** may be resolved into two orthogonal components, $\mathbf{E}_\sigma$ and $\mathbf{E}_\pi$, where $\mathbf{E}_\sigma$ is perpendicular to both the line NO and the scattering plane NOP, and $\mathbf{E}_\pi$ is the component parallel to this plane. Because of the random nature of the direction of **E**, the mean square values are equal: $\overline{E_\sigma^2} = \overline{E_\pi^2} = \overline{E^2}$. The intensity is therefore evenly divided between the two polarizations:

$$I_{0\sigma} = I_{0\pi} = I_0/2. \tag{7.20}$$

The intensity scattered to point P is the sum of the intensities for the two polarizations. For σ polarization, $\alpha = 90°$, but for π polarization, $\alpha = 90° - 2\theta$, where θ is the scattering angle. Therefore, the intensity scattered to the point of observation P is

$$I_P = \frac{I_0}{2}\left(\frac{\mu_0}{4\pi}\right)^2\left(\frac{q^4}{m^2r^2}\right)\left[1+\cos^2(2\theta)\right]. \tag{7.21}$$

This is the Thompson equation for the scattering of an unpolarized x-ray beam by a single electron. In an XRD experiment, all of the terms in this equation are constant except for $[1 + \cos^2(2\theta)]$, which is known as the polarization factor.

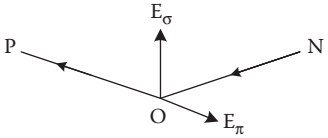


FIGURE 7.5
Scattering of a randomly polarized x-ray beam by an electron.

7.2.2.2 Scattering of X-Rays by an Atom

Each electron in an atom scatters an x-ray beam according to the Thompson equation. The net effect of the electrons is described by the atomic scattering factor *f*, defined as the ratio of the amplitude of the wave scattered by the atom to the amplitude that would be scattered by one electron. The atomic scattering factor depends on the atomic number, scattering angle, and x-ray wavelength. The wavelength dependence arises from the scattering angle dependence and also from the anomalous dispersion corrections. Anomalous dispersion corrections are significant when the incident x-ray wavelength is comparable with the absorption edge of the scatterer. The imaginary component of the anomalous dispersion correction accounts for x-ray absorption.

Numerical values for the atomic scattering factors may be obtained using analytic expressions available in the International Tables for X-Ray Crystallography.[4] These expressions are best fits to experimentally determined atomic scattering factors, and have the form

$$f_0(x,\text{atom}) = \sum_{i=1}^{4}\left\{a(\text{atom},i)\exp\left[-b(\text{atom},i)x^2\right]\right\} + c(\text{atom}), \tag{7.22}$$

where $x = \sin\theta_B/\lambda$. The nine coefficients, a(atom,i), i = 1,2,3,4, b(atom,i), i = 1,2,3,4, and c(atom), are given in the International Tables for X-Ray Crystallography for many elements. Figure 7.6 shows the atomic scattering factors for the atoms N, C, Si, Ga, In, As, and P. In the case of forward scattering ($\theta = 0$), the atomic scattering factor is equal to the atomic number. There is a monotonic decrease in the atomic scattering factor with scattering angle, which diminishes the intensities of reflections with large Bragg angles.

Anomalous dispersion corrections to the atomic scattering factors are also provided in the International Tables for X-Ray Crystallography. These corrections account for the fact that bound electrons in the atom scatter differently from free electrons, if the frequency of the incident radiation is comparable with an absorption frequency of the atom. The anomalous dispersion corrections are complex, to account for corrections to the magnitude

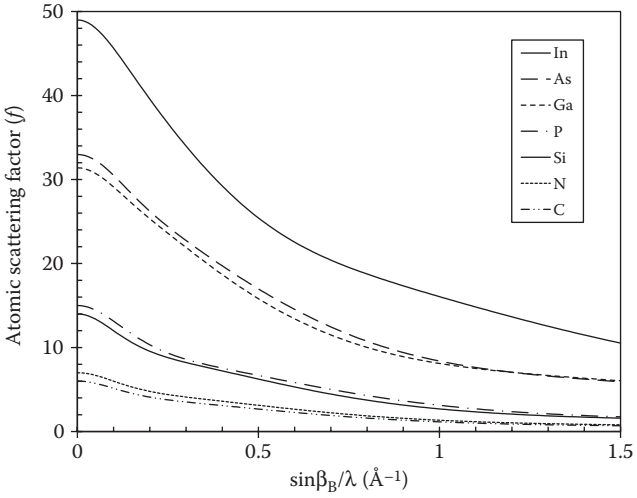


FIGURE 7.6
Atomic scattering factors for the atoms N, C, Si, Ga, In, As, and P.

and phase of the scattered radiation. For a particular atom scattering a wavelength λ, the atomic scattering factor may be adjusted for anomalous dispersion by

$$f(x,\text{atom},\lambda) = f_0(x,\text{atom}) + f_{\text{corr}}(\text{atom},\lambda), \tag{7.23}$$

where f_{corr}(atom, λ) is the anomalous dispersion correction, which has been tabulated for a number of commonly used x-ray wavelengths.

7.2.2.3 Scattering of X-Rays by a Unit Cell

The x-ray amplitude scattered by a unit cell is the vector sum of the amplitudes scattered by the individual atoms, taking into account phase differences. It is described by the structure factor, the magnitude of which is normalized to the scattering amplitude for a single electron. If the unit cell contains N atoms with atomic scattering factors f_n, then the structure factor for the *hkl* reflection is given by

$$F_{hkl} = \sum_{n=1}^{N} f_n \exp\{2\pi i(hu_n + kv_n + lw_n)\}, \tag{7.24}$$

where (u_n, v_n, w_n) is the position of the nth atom normalized to the primitive unit cell vectors **a**, **b**, and **c**.

7.2.2.3.1 Structure Factor in Diamond and Zinc Blende Materials

For the zinc blende crystal structure made of atoms A and B, the structure factor may be simplified to the following when the indices are not mixed (i.e., either all even or all odd):

$$F_{hkl} = 4\{f_A + f_B \exp[(i\pi/2)(h+k+l)]\}. \tag{7.25}$$

The intensity of a reflection is proportional to $|F|^2$, which is found by multiplying the structure factor by its complex conjugate. For zinc blende semiconductors, if the anomalous dispersion is neglected, there are four cases of $|F|^2$, which are

$$\begin{aligned}
&|F|^2 = 0, && \text{when } h,k, \text{ and } l \text{ are mixed even and odd (forbidden reflection);}\\
&|F|^2 = 16(f_A - f_B)^2, && \text{when } (h+k+l) \text{ is an odd multiple of 2}\\
&&& \text{(weak in zinc blende, forbidden in diamond);}\\
&|F|^2 = 16(f_A{}^2 + f_B{}^2), && \text{when } (h+k+l) \text{ is odd (strong);}\\
&|F|^2 = 16(f_A + f_B)^2, && \text{when } (h+k+l) \text{ is an even multiple of 2 (very strong).}
\end{aligned} \tag{7.26}$$

If anomalous dispersion is included, Equation 7.24 must be applied, using the values of $f(x,$ atom, $\lambda)$ determined in Section 7.2.2.2. An immediate consequence is that $F(hkl) \neq F(\bar{h}\,\bar{k}\,\bar{l})$.

An important result from Equation 7.24 is that those *hkl* reflections with mixed even and odd indices are forbidden for the zinc blende and diamond crystals. The 112, 001, and

003 are examples of these forbidden reflections. The 002, 006, 222, and 024 are examples of reflections that are forbidden in the diamond structure and weak in zinc blende crystals. The 113, 115, and 333 are strong, and the 004, 026, and 044 are very strong in both types of crystals.

For (001) heteroepitaxy of zinc blende semiconductors, the 002, 004, and 006 reflections may be excited from planes parallel to the interface (symmetric reflections). Of these, the 004 is usually preferred because of its greater intensity.

7.2.2.3.2 Structure Factor in Wurtzite Materials[5]

For a binary III-nitride semiconductor (GaN, AlN, InN) with the hexagonal wurtzite structure, the unit cell contains group III atoms at (0, 0, 0) and (1/3, 2/3, 1/2) and nitrogen atoms at (0, 0, 3/8) and (1/3, 2/3, 1/2). The structure factor for the *hkil* reflection (also referred to as the *hkl* reflection, because i is not independent but is equal to $-h-k$) is therefore given by

$$F_{hkl}=\left(\begin{array}{l} f_{\mathrm{III}}\left\{1+\exp\left[i2\pi\left(\dfrac{h+2k}{3}+\dfrac{1}{2}\right)\right]\right\} \\ +f_N\left\{\exp\left[i2\pi\left(\dfrac{31}{8}\right)\right]+\exp\left[i2\pi\left(\dfrac{h+2k}{3}\right)+\dfrac{71}{8}\right]\right\} \end{array}\right), \tag{7.27}$$

where f_{III} is atomic scattering factor for the anion (Ga, In, or Al) and f_N is the atomic scattering factor for nitrogen.

For a ternary alloy such as $Al_xGa_{1-x}N$, the structure factor depends on the aluminum mole fraction. In the absence of ordering, the anion sites are randomly occupied by Al or Ga, so the effective atomic scattering factor for the column III sublattice is given by

$$f_{\mathrm{III}}=xf_{Al}+(1-x)f_{\mathrm{Ga}}, \tag{7.28}$$

but the structure factor may be determined as before using this value. For $In_xGa_{1-x}N$, however, there is the possibility of ordering on the column III sublattice, which alters the structure factor because the effective atomic scattering factors for the (0, 0, 0) and (1/3, 2/3, 1/2) sites will be different. If these effective atomic scattering factors are f_1 and f_2, respectively, then the structure factor is given by

$$F_{hkl}=\left(\begin{array}{l} f_1+f_2\exp\left[i2\pi\left(\dfrac{h+2k}{3}+\dfrac{1}{2}\right)\right] \\ +f_N\left\{\exp\left[i2\pi\left(\dfrac{31}{8}\right)\right]+\exp\left[i2\pi\left(\dfrac{h+2k}{3}\right)+\dfrac{71}{8}\right]\right\} \end{array}\right). \tag{7.29}$$

If the fraction of indium on (1/3, 2/3, 1/2) sites is r_{In}, then $r_{\mathrm{In}}=1$ corresponds to complete ordering while $r_{\mathrm{In}}=0.5$ corresponds to random alloying. In terms of this unitless parameter, the structure factor for a partially ordered alloy of $In_xGa_{1-x}N$ is given by

$$F_{hkl} = \begin{pmatrix} 2x(1-r_{\mathrm{In}})f_{\mathrm{In}} + \left[1-2x(1-r_{\mathrm{In}})\right]f_{\mathrm{Ga}} \\ +\left[2r_{\mathrm{In}}xf_{\mathrm{In}} + (1-2r_{\mathrm{In}}x)f_{\mathrm{Ga}}\right]\exp\left[i2\pi\left(\frac{h+2k}{3}+\frac{1}{2}\right)\right] \\ +f_N\left\{\exp\left[i2\pi\left(\frac{31}{8}\right)\right]+\exp\left[i2\pi\left(\frac{h+2k}{3}\right)+\frac{71}{8}\right]\right\} \end{pmatrix}. \tag{7.30}$$

As a consequence, the 001 reflection, which is forbidden in an $In_xGa_{1-x}N$ alloy with random occupation of the column III sublattice, takes on a nonzero structure factor in the presence of ordering. The observation of this reflection therefore indicates the presence of ordering, and the ordering may be quantized by comparing the intensity of this reflection with that of the 004 reflection, which is unaffected by ordering.

7.2.2.4 Intensities of Diffraction Profiles

The absolute intensities of diffraction peaks depend on many factors and are difficult to predict. However, the relative intensities for *hkl* reflections from single perfect crystals using an ideal diffractometer with zero divergence and monochromatic radiation can be estimated by[6]

$$I \propto \frac{|F|^2\left(1+\cos^2\theta_B\right)e^{-2M}}{\sin\theta_B\cos\theta_B}, \tag{7.31}$$

where $|F|$ is the magnitude of the structure factor, $(1 + \cos^2 \theta_B)/(\sin \theta_B \cos \theta_B)$ is the Lorentz polarization factor, and e^{-2M} is the temperature factor. As an example, for a perfect GaAs(001) crystal, this equation predicts that the 004 reflection will have about 250 times the peak intensity of the 002 reflection, for the case of Cu kα radiation. This is approximately the intensity ratio observed using a double-crystal diffractometer or Bartels diffractometer.

It should be noted that for the case of a real diffractometer, the intensities of broader diffraction lines are enhanced; this is because a broader Bragg peak can reflect a greater portion of the beam that is both divergent and contains a spread of wavelengths.

Relative intensities of lines are of more fundamental importance than the absolute intensities, which depend strongly on the instrument. Therefore, the temperature factor, which scales all reflections in equal fashion, is of little importance unless measurements are taken at different temperatures. The temperature factor has been treated in detail by James.[7]

7.2.3 Dynamical Diffraction Theory

XRD profiles (or rocking curves) from heteroepitaxial structures often exhibit interesting shapes and multiple peaks, which are difficult to interpret directly. Instead, the depth profiles of strain and composition are guessed based on the growth process. Based on this guess, the XRD profile is simulated and then compared with the experimental results. Subsequent refinement of the model structure continues until there is reasonable agreement between the simulated and experimental profiles. Then, the simulation structure is assumed to closely represent the physical sample.

Early work was often based on the kinematical theory, which has been described by Speriosu and coworkers.[8–13] This theory, appropriate for thin heteroepitaxial films as well as powders and polycrystalline samples, treats volume elements of the crystal independently, except for the inclusion of incoherent power losses to the diffracted beam. Important applications of the kinematical theory include the simulation of XRD profiles from semiconductor multilayers and superlattices, thin films, and ion-implanted regions of crystals. The advantage of the kinematical treatment is that it reduces computational complexity and time compared with the more generally applicable dynamical theory. However, this advantage has become less important due to the advances in computer speed.

The dynamical theory has been described in detail by Darwin,[14] Ewald,[15–17] von Laue,[18] Batterman and Cole,[19] Zachariasen,[20] Klar and Rustichelli,[21] Takagi,[22,23] and Taupin.[24] In this theoretical treatment, all wave interactions within the crystal are included. The dynamical theory must be applied for the accurate determination of diffraction profiles from thick or nearly perfect crystals, but is generally applicable to heteroepitaxial structures as well.

The foundation of the dynamical theory is a solution of Maxwell's equations in the periodic electron density of the crystal. It has enabled the calculation of the intensities and shapes of diffraction profiles from infinitely thick, perfect crystals, which serves as a starting point for the study of real crystals. Dynamical theory has also been extended to the case of laminar (layered) structures[21,22,25–27] and crystals with any arbitrary distortion,[23,24] for the simulation of diffraction profiles from heteroepitaxial multilayered structures, and this aspect is described in Section 7.12. In addition, recent work[28–32] has extended the dynamical diffraction model to include the effects of dislocations on the lattice, and this is described in detail in Section 7.12.1.

7.2.3.1 Intrinsic Diffraction Profiles for Perfect Crystals

Usually, the substrate for a heteroepitaxial structure may be treated approximately as a perfect crystal. In addition, the diffraction profile for a perfect crystal serves as a starting point for the analysis of an imperfect heteroepitaxial layer.

The solution of Maxwell's equations in the crystal yields the Takagi–Taupin equations,[22–24] which describe the change in scattering amplitude with depth in the diffracting crystal. The complex scattering amplitude is the ratio of the diffracted and incident waves, which exchange energy through multiple scattering. Taupin[24] combined this set of equations into a single differential equation for the centrosymmetric Bragg case, which was subsequently generalized to the case of polar crystals by Bartels.[27] The resulting equation is

$$-i\frac{dX}{dT} = X^2 - 2\eta X + 1, \tag{7.32}$$

where T is the thickness parameter, given by

$$T = t\frac{\pi\Gamma\sqrt{F_H F_{\bar{H}}}}{\lambda\sqrt{|\gamma_0\gamma_H|}}, \tag{7.33}$$

where:

- t is the depth measured from the diffracting surface
- η is the deviation parameter
- X is the scattering amplitude, given by

$$X = \sqrt{\frac{F_{\bar{H}}}{F_H}}\sqrt{\frac{\gamma_0}{\gamma_H}}\frac{D_H}{D_0}. \tag{7.34}$$

D_H is the amplitude of the diffracted wave
D_0 is the amplitude of the incident wave
F_H is the structure factors for the hkl
$F_{\overline{H}}$ is the structure factors for the $\overline{h}\,\overline{k}\,\overline{l}$ reflection

γ_H and γ_0 are the direction cosines of the diffracted and incident waves with respect to the inward surface normal, and their ratio is the asymmetry factor b,

$$b = \gamma_0 / \gamma_H, \tag{7.35}$$

which accounts for any differences in the angles of incidence and exit for the x-rays. For the Bragg case, the diffracted beam passes back through the same surface through which the incident beam enters, and $b < 0$. For the symmetric Bragg case, that is, equal angles of incidence and exit, $b = -1$.

The deviation parameter η describes the departure from the Bragg condition,

$$\eta = \frac{-b(\theta - \theta_B)\sin(2\theta_B) - 0.5(1-b)\Gamma F_0}{\sqrt{|b|}\Gamma\sqrt{F_H F_{\overline{H}}}}, \tag{7.36}$$

where:
θ_B is the Bragg angle
θ is the actual angle of incidence on the diffracting planes
Γ is given by

$$\Gamma = \frac{r_e \lambda^2}{\pi V}, \tag{7.37}$$

where:
r_e is the classical electron radius, 2.818×10^{-5} Å
λ is the x-ray wavelength
V is the crystal volume for which we have calculated the structure factor (For a cubic crystal, $V = a_0{}^3$.)

When the polarization of the incident beam is in the plane of incidence (π polarization), $C_\pi = |\cos(2\theta_B)|$, and when the x-rays are polarized perpendicular to the plane of incidence (σ polarization), $C_\sigma = 1$.

For an infinitely thick, perfect crystal, the solution of the Takagi–Taupin equations yields the Darwin–Prins formula:[33]

$$X = \eta \pm \sqrt{\eta^2 - 1}. \tag{7.38}$$

Here, the sign must be chosen to be opposite to that of the real component of η; in other words,

$$X = \eta - \mathrm{Sign}(\mathrm{Re}(\eta))\sqrt{\eta^2 - 1}. \tag{7.39}$$

This equation allows the calculation of the diffracted intensity $I = |X|^2$ as a function of angle θ for a perfect crystal. Figure 7.7 shows the calculated 004 rocking curve for an infinitely thick, perfect GaAs(001) crystal. Since the thickness of a substrate is typically many times the extinction depth, Equation 7.33 can be used to calculate the diffraction from the substrate, and this computation is the first step to calculating diffraction from a multilayer heterostructure.

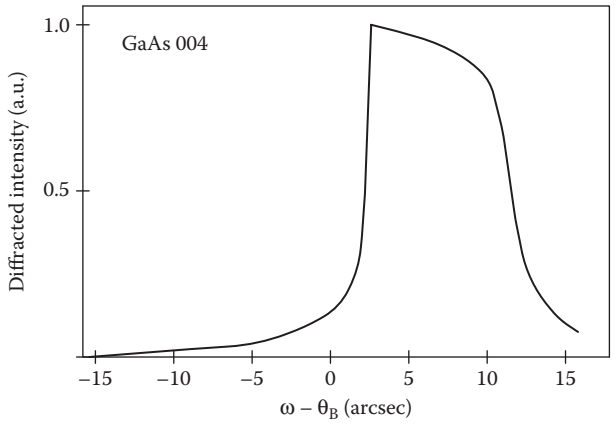


FIGURE 7.7
004 rocking curve for a perfect GaAs(001) crystal, calculated using the Darwin–Prins formula.

7.2.3.2 Intrinsic Widths of Diffraction Profiles

The Darwin–Prins equation predicts that intrinsic diffraction profiles for perfect crystals should have finite width. Physically, this arises because the incident x-ray intensity reduces with each successive plane due to extinction and absorption, and the destructive interference with $\theta \neq \theta_B$ is not perfect. We thus expect stronger reflections to have broader intrinsic profiles.

In the absence of absorption, the intrinsic diffraction profile has a flat top corresponding to total Bragg reflection over a finite angular range W_S. This width approximates the full width at half maximum (FWHM) for the intrinsic profile and can be calculated as

$$W_S = \frac{2r_e\lambda^2 C|F_H|}{\pi V \sin(2\theta_B)}, \tag{7.40}$$

where:

- W_S is the natural width of the diffraction profile for the symmetric case (equal angles of incidence and exit)
- r_e is the classical electron radius, 2.818×10^{-5} Å
- λ is the x-ray wavelength
- V is the crystal volume for which we have calculated the structure factor $V = a_0^3$ for a cubic crystal
- C is the polarization factor, which is usually assumed to be 1
- $|F_H|$ is the magnitude of the structure factor for the *hkl* reflection
- θ_B is the Bragg angle

The FWHMs for symmetric reflections from absorbing crystals are closely approximately by Equation 7.40 except for the very weak reflections.

For the case of an asymmetric reflection (Bragg planes inclined to the crystal surface), the natural width of the diffraction profile depends on the angle of incidence for the exploring beam. If the (*hkl*) planes are inclined to the (*mno*) surface by an angle Φ, then $(\theta_B + \Phi)$ incidence will give a narrower rocking curve than $(\theta_B - \Phi)$ incidence. If W_S is the symmetric profile width for $\Phi = 0°$, then

$$W_+ = W_S/\sqrt{b};(\theta_B + \Phi) \text{ incidence,}$$

and

$$W_- = W_S\sqrt{b};(\theta_B - \Phi) \text{ incidence.} \tag{7.41}$$

This effect is significant for highly asymmetric reflections. For example, in the case of the 353 Bragg reflection from (001) GaAs using Cu kα radiation, $b = -0.0128$ and $W_-/W_+ = 78$. The tables in Appendix VII provide the natural widths of rocking curves for selected semiconductor crystals, for the case of Cu kα radiation.

7.2.3.3 Extinction Depth and Absorption Depth

The intensity of an exploring x-ray beam diminishes with depth in a diffracting crystal, as intensity is transferred to the diffracted beam; this effect is called "extinction." On top of this, the exploring beam loses energy to photoelectric absorption within the crystal. Because of these two effects, the x-ray beam probes a finite depth of a crystal specimen. In the absence of absorption, most of the integrated intensity for a reflection originates within a distance t_{ext} from the surface, which is called the "extinction depth." It is given by[20,34]

$$t_{ext} = \frac{\pi\lambda\sqrt{|\gamma_0\gamma_H|}}{\Gamma|F_{\bar{H}}|}. \tag{7.42}$$

Here $F_{\bar{H}}$, which is the structure factor for the $\bar{h}\,\bar{k}\,\bar{l}$ reflection, is due to the real components of the atomic scattering factors. Typical extinction depths for reflections used in the characterization of heteroepitaxial layers are of the order of 10 μm. Weak reflections may have extinction depths of hundreds of microns; an example is the 006 reflection from GaAs(001), for which t_{ext} = 531 μm using Cu kα radiation. Strong reflections have smaller extinction depths; an example is the 224 reflection from HgTe(001), for which t_{ext} = 1.6 μm using Cu kα radiation.

Experimentally, diffraction profiles from layers of thickness less than $t_{ext}/100$ are obtained with great effort due to their weak intensity. Glancing angle geometry[35] is therefore sometimes used for thin layers to reduce the extinction depth. For typical double-axis XRD experiments with zinc blende heteroepitaxial layers, 0.1 μm is an approximate minimum thickness required to obtain usable intensity.

If the extinction is weak, then the thickness of crystal that contributes most of the diffracted intensity is equal to the absorption depth t_{abs}, given by[20,34]

$$t_{abs} = \frac{\lambda}{2\pi\,\Gamma\left|F_0'\right|}\frac{\left|\gamma_0\gamma_H\right|}{\left(\left|\gamma_0\right|+\left|\gamma_H\right|\right)}, \tag{7.43}$$

where $|F_0'|$ is the magnitude of the 000 structure factor that is due to the imaginary components of the atomic scattering factors. Typical absorption depths in semiconductors are of the order of 10 μm, but depend strongly on the crystal density. Absorption depths with the 004 reflection, and Cu kα radiation, are 8.0 μm in GaAs (density = 5.32 g cm^{-3}) and 1.3 μm in CdTe (density = 8.17 g cm^{-3}).

A practical implication of the absorption depth is the following: t_{abs} is (approximately) the maximum thickness of a mismatched heteroepitaxial layer through which strong substrate diffraction may be observed.

In general, both extinction and absorption may be important, and the penetration depth t_p for the x-ray beam is then

$$t_p = \left(\frac{1}{t_{ext}} + \frac{1}{t_{abs}} \right)^{-1}. \tag{7.44}$$

Values of the extinction, absorption, and penetration depth are tabulated in Appendix VII.

7.2.4 X-Ray Diffractometers

The high-resolution x-ray diffractometers used for the characterization of heteroepitaxial semiconductors are usually of the double-axis or triple-axis type. Therefore, there will be two or three axes perpendicular to the *plane of the diffractometer*.*

The source of x-rays is usually a sealed x-ray tube, the output of which consists of strong characteristic lines superimposed on a broad spectrum (the "braking radiation" or Bremsstrahlung). Typically, the kα lines of Co, Cr, Cu, Fe, or Mo are used. The kα spectra of these elements contain one very strong peak, the $k\alpha_1$, and a strong peak, the $k\alpha_2$. The wavelengths for these lines are given in Table 7.1 for the commonly used anodes. The $k\alpha_1$ typically has twice the integrated intensity of the $k\alpha_2$, and the two wavelengths usually differ by 0.2%–0.6% in wavelength, with the $k\alpha_2$ at the longer wavelength.

Figure 7.8 shows the Cu kα spectrum.[39] The Cu $k\alpha_1$ has peak intensity at $\lambda_1 = 1.540594$ Å[37] and an FWHM of $W_1 = 4.61(9) \times 10^{-4}$ Å.[40] The Cu $k\alpha_2$ has peak intensity at $\lambda_2 = 1.544423$ Å[37] and an FWHM of $W_2 = 6.1(4) \times 10^{-4}$ Å.[40] Both peaks are asymmetric Lorentzian distributions. For a typical 2 kW Cu *k*α source (sealed tube, fixed anode), the available intensity is ~10^6 photons s^{-1} and the background intensity is <0.1 photon/s, providing a dynamic range of seven orders of magnitude.[41]

The output intensity for the x-ray tube increases linearly with the emission current and sublinearly with the accelerating voltage. Sometimes, rotating anode tubes are used because they permit higher operating currents and therefore intensity. Occasionally, synchrotron radiation is used if very high intensity is needed (i.e., for the characterization of very thin layers).

TABLE 7.1

Commonly Used kα Wavelengths

Element	$k\alpha_1$ (Å)	$k\alpha_2$ (Å)	kα (Weighted Average) (Å)
Co	1.788965[36]	1.79285[36]	1.790260[a]
Cr	2.28970[36]	2.293606[36]	2.29100[a]
Cu	1.540594[37]	1.544423[37]	1.541942[38]
Fe	1.936042[36]	1.939980[36]	1.937355[a]
Mo	0.70930[36]	0.713590[36]	0.710730[a]

[a] The $k\alpha_1$ was assigned twice the weight of the $k\alpha_2$.

* This is the plane that includes the beams incident on and diffracted by the specimen.

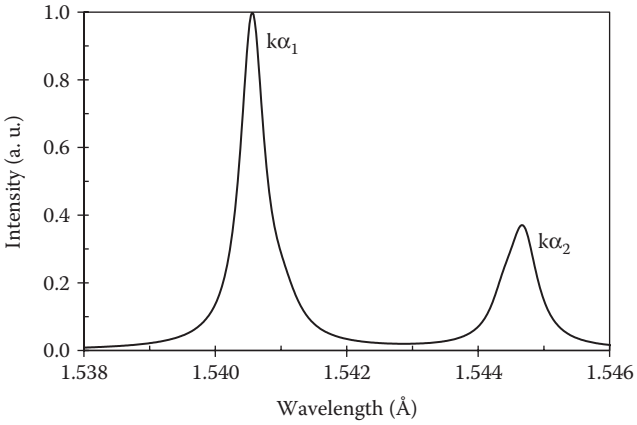


FIGURE 7.8
Cu kα spectrum.

The raw x-ray beam produced by an x-ray tube is divergent and contains a spread of wavelengths. In order to reduce these effects, the first axis of the diffractometer is fitted with one or more diffracting crystals, in order to condition the x-ray beam.

The specimen is mounted on the second axis, which is the most critical in the instrument design. For high-resolution measurements, the axis 2 goniometer must have a step size of less than 1 arcsec, with 0.1 arcsec being typical. Usually, only peak separations need to be measured, and these can be found from the step size and the number of steps that have been taken. Sometimes, absolute angle encoders are affixed to the second axis of the diffractometer.

Other axes are provided that allow for tilt adjustment of the monochromator and the specimen. These make it possible to bring the diffraction vectors of the monochromator crystal(s) and specimen into the plane of the diffractometer. Usually, another axis is provided for the rotation of the specimen about its azimuth (about the surface normal). Axis 2 is driven by a stepper motor to facilitate computer control. Often, the other axes are also computer controlled so that series of experiments may be fully automated.

Usually, a scintillation detector is used. This device uses a NaI crystal followed by a photomultiplier. Each x-ray photon produces a pulse of current from the scintillation detector, and these pulses are counted for the measurement of the intensity. However, the detector will saturate when the current pulses begin to overlap, typically at ~10^6 counts per second.

The following sections give some specifics of the three most important instruments for characterizing heteroepitaxial layers: the double-crystal diffractometer, the Bartels diffractometer, and the triple-axis diffractometer.

7.2.4.1 Double-Crystal Diffractometer

The double-crystal diffractometer is a double-axis instrument in which matched crystals are placed on the first and second axes. As shown in Figure 7.9, there are two possible configurations for the instrument. In the (+,–) configuration, the first crystal and the specimen bend the beam in opposite directions (counterclockwise and clockwise, respectively), but in the (+,+) configuration, both crystals bend the beam in the same direction. Only the (+,–) configuration is useful because the (+,+) setup is dispersive. This is because the first

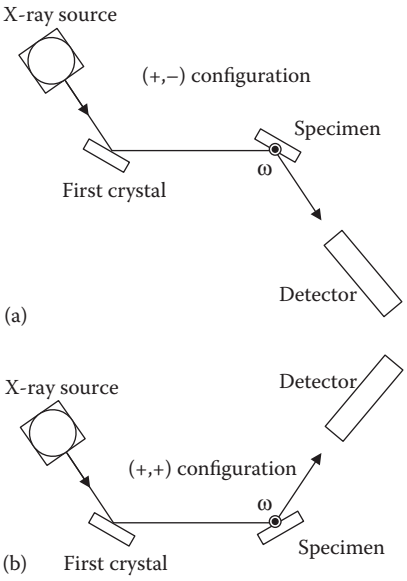


FIGURE 7.9
Double-crystal diffractometer.

crystal does not act as a true monochromator, but instead disperses the various wavelengths according to the Bragg law.

This can be understood with the aid of the Dumond diagram[42] shown in Figure 7.10. The dispersion relationship for the (fixed) first crystal is given by*

$$\lambda = 2d_1 \sin\theta, \tag{7.45}$$

where d_1 is the spacing of the diffracting planes in the first crystal. With the (+,–) parallel configuration, the dispersion relationship for the specimen crystal is

$$\lambda = 2d_2 \sin(\theta + \omega), \tag{7.46}$$

where:

d_2 is the spacing of the diffracting planes in the specimen
ω is the rocking angle (the rotation of the specimen crystal from the position with peak intensity)

Each crystal has a finite rocking curve width, as shown in the diagram. The diffracted intensity versus angle may be found by integrating the overlap of the two dispersion functions for each angle. In the (+,–) parallel arrangement, significant overlap of the two dispersion characteristics occurs only with $\omega \approx 0$. But for the (+,+) antiparallel configuration shown in Figure 7.10b, the dispersion relationship for the specimen is given by

$$\lambda = 2d_2 \sin(-\theta + \omega), \tag{7.47}$$

* A first-order reflection was assumed ($n = 1$).

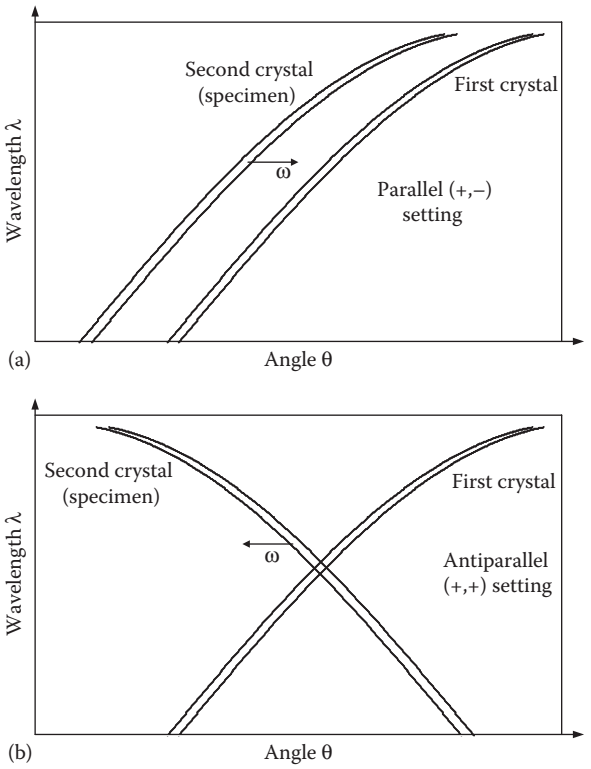


FIGURE 7.10
Dumond diagram for double-crystal diffractometer with matched crystals. Each crystal has a dispersion relation given by the Bragg equation: $\lambda = 2d \sin\theta/n$. The finite angular width for each crystal (exaggerated here) is due to the finite rocking curve width for the crystal. (a) In the parallel (+,–) arrangement, the rocking curve is narrow, because the two dispersion relationships overlap only for a narrow range of ω. (b) In the antiparallel (+,+) setting, a broad rocking curve is obtained because of the wide overlap of the two dispersion characteristics.

and it can be seen that the two dispersion characteristics will overlap for a significant range of ω if the x-ray source produces a range of wavelengths.

Mathematically, the dispersion of the double-crystal diffractometer (the broadening due to the source wavelength spread) is given by

$$\frac{\Delta\omega}{\Delta\lambda} = \frac{\tan\theta_1 \pm \tan\theta_2}{\lambda}, \tag{7.48}$$

where:

- θ_1 is the Bragg angle for the first crystal
- θ_2 is the Bragg angle for the specimen
- $\pm$ plus applies to the parallel arrangement while minus applies to the antiparallel arrangement

The double-crystal diffractometer is nondispersive only if the first crystal and specimen are matched ($\theta_1 = \theta_2$) and the (+,–) parallel configuration is used. A practical difficulty in the use of the double-crystal instrument with a heteroepitaxial sample is that the first crystal cannot be matched to both the substrate and the heteroepitaxial layer. A second limitation

is that measurements cannot be taken in the (+,+) configuration with high resolution; this precludes the use of the double-crystal diffractometer for absolute measurements of lattice constants by the Bond method.

7.2.4.2 Bartels Double-Axis Diffractometer

The Bartels diffractometer uses an arrangement of two channel-cut crystals with four diffracting surfaces as a monochromator,* as shown in Figure 7.11. Typically, the monochromator uses four symmetric 220 or 440 reflections from two channel-cut Ge crystals with (110) faces. Each of the channel-cut crystals acts as a double-crystal diffractometer in the (+,–) configuration. In the first channel-cut crystal, the first reflection passes a wide range of wavelengths, but each wavelength is diffracted at a particular angle. The second reflection accepts this entire wavelength spread, but bends the beam back into line with the source beam. The third reflection (from the first surface of the second channel-cut crystal) can accept a narrow piece of this spectrum, because this crystal is antiparallel with the second and its acceptance angle for a particular wavelength is approximately the intrinsic rocking curve width for this reflection. The fourth reflection brings the beam back into the line of the source beam. Therefore, the Bartels monochromator produces a conditioned beam with a divergence and wavelength spread that are both determined by the intrinsic rocking curve width of the monochromator reflections. Using Ge 440 reflections, the conditioned beam exiting the monochromator has a divergence of 5 arcsec and a wavelength spread of $\Delta\lambda/\lambda = 2.3 \times 10^{-5}$ (23 parts per million). Because of this, diffraction profiles may be measured in either geometry, (+,–,–,+,–) or (+,–,–,+,+), with very little dispersion. Also,

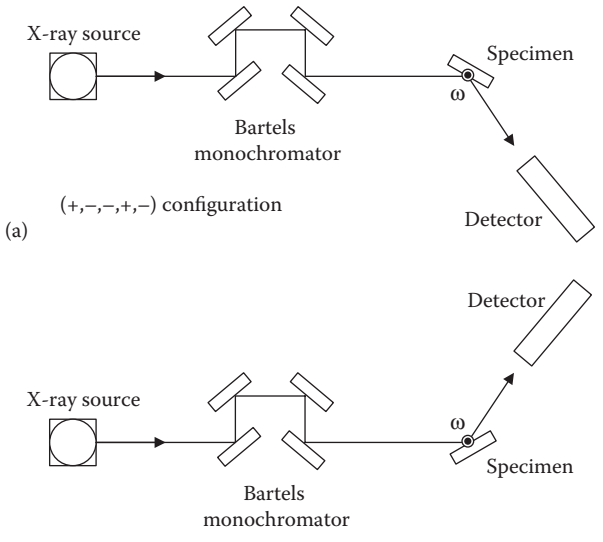


FIGURE 7.11
Bartels diffractometer. (a) (+,–,+,–,+) configuration. (b) (+,–,–,+,+) configuration.

* Sometimes this arrangement is called a "monochrocollimator" because it reduces the beam divergence as well as the wavelength spread.

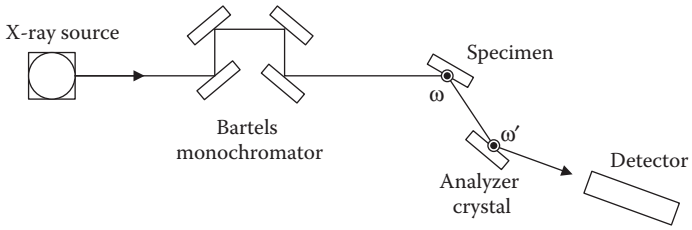


FIGURE 7.12
Triple-axis diffractometer.

the monochromator need not be matched to the specimen. Therefore, a single monochromator can be used to measure near-ideal rocking curves for a wide range of specimen crystals, even specimens that contain layers with different lattice constants.

7.2.4.3 Triple-Axis Diffractometer

The triple-axis diffractometer proposed by Fewster[43,44] has additional versatility due to the use of an analyzer placed between the specimen and the detector. This instrument is illustrated in Figure 7.12. The analyzer is a crystal oriented to diffract the beam from the specimen, with an angle of acceptance equal to its intrinsic rocking curve width. Typically, a single analyzer reflection is used; however, if the analyzer is a channel-cut crystal, then it may be arranged to diffract the beam two or more times. Often, a Bartels four-reflection arrangement is used for the analyzer. The essential behavior of the instrument is preserved in all of these cases, but very high resolution is obtained with a Bartels analyzer.

There are two modes of operation for the triple-axis diffractometer. These are the diffraction profile mode and the mapping mode.

In the diffraction profile mode, the computer control is set to couple the rotations of the specimen, ω, and the analyzer crystal, ω′, such that ω′ = 2ω. The final profile is given by

$$R(\omega) = \int_{\omega'}\int_{\alpha} R_m(\alpha) R_s(\omega - \alpha) R_a(2\omega - \omega')\, d\alpha\, d\omega', \tag{7.49}$$

where $R_m(\phi)$, $R_s(\phi)$, and $R_a(\phi)$ are the reflectivity profiles for the monochromator, specimen, and analyzer, respectively. This ω − 2ω scan yields a near-ideal rocking curve for the specimen. It is similar to the ω scan measured by the Bartels diffractometer, with two differences: (1) in the triple-axis case, the rocking curve is relatively unaffected by sample curvature and distortions, and (2) for metamorphic structures containing dislocations, the ω scan measured by the Bartels diffractometer is sensitive to the angular mosaic spread, as well as the d-spacing mosaic spread introduced by dislocations, but the triple-axis ω − 2ω scan is only sensitive to the d-spacing mosaic spread of dislocations. This difference must be considered when using the rocking curve width to estimate the dislocation density.

7.2.5 Resolution of X-Ray Diffraction Measurements and the Effect of Finite Counting Statistics

Understanding the resolution limits of x-ray rocking curve measurements is important when studying the strain in heteroepitaxial layers and, in particular, the initial onset of lattice relaxation and the apparent critical layer thickness (CLT). Often, researchers estimate

the resolution as a single number or give a unitless constant of proportionality (r) to relate the experimental resolution to the diffraction peak widths. Kujofsa and Ayers[45] showed that the resolution of x-ray rocking measurements also depends on the effective number of counts (x-ray photons) observed, as well as the diffraction peak FWHM and the Bragg angle. In this work, Gaussian and Lorentzian profiles were considered, and the resolution was related to the effective counts and peak width for both cases. This work is summarized below, but a more detailed treatment can be found in Appendix VIII.

In this work, Gaussian or Lorentzian profiles were considered in the presence of statistical noise, and it was shown that the uncertainty in the peak position $\Delta\theta$ is proportional to the diffraction peak FWHM and is given by

$$\Delta\theta = C_P\beta, \tag{7.50}$$

where:

C_P is a constant of proportionality that is dependent on the signal-to-noise ratio, and therefore the effective number of counts

β is the FWHM for the diffraction profile

In the case of a Gaussian rocking curve profile, the constant of proportionality is given by

$$C_P = \frac{1}{2\sqrt{\ln(2)}}\sqrt{-\ln\left(1-\frac{1}{\sqrt{N_{\text{Eff}}}}\right)} \text{ (Gaussian profile)}, \tag{7.51}$$

where N_{Eff} is the effective number of counts, whereas in the case of a Lorentzian distribution, the constant of proportionality is

$$C_P = \frac{1}{2}\sqrt{\frac{1}{\sqrt{N_{\text{Eff}}}-1}} \text{ (Lorentzian profile)}. \tag{7.52}$$

The additive noise also introduces an error in the measurement of the peak width, as well as its position. It was shown that this uncertainty ($\Delta\beta$) is also proportional to the peak width, but with a different constant of proportionality,

$$\Delta\beta = C_w\beta\,. \tag{7.53}$$

The constant of proportionality is given by

$$C_w = \frac{1}{\sqrt{N_{\text{Eff}}\cdot\ln(2)}} \text{ (Gaussian profile)} \tag{7.54}$$

and

$$C_w = \frac{1}{\sqrt{N_{\text{Eff}}}} \text{ (Lorentzian profile)} \tag{7.55}$$

for the Gaussian and Lorentzian cases, respectively. The above derivation is important when considering strain resolution from peak position and peak width measurements; in the case when the lattice is tetragonally distorted and a symmetric reflection is considered, the strain resolution corresponding to the minimum resolvable peak shift is given by

$$R_P(\text{strain}) = \frac{2\nu}{1-\nu} C_P \beta \cot(\theta_B), \tag{7.56}$$

where ν is the Poisson ratio. Similarly, in the case of peak width measurements in the presence of dislocations, the strain resolution corresponding to the minimum detectable increase in the peak width by threading dislocations and the associated misfit dislocations is

$$R_W(FWHM) = \frac{1}{\sqrt{4.36x}} C_W \beta \cos\alpha \cos\phi, \tag{7.57}$$

where:

- α is the angle between the Burgers vector and line vector for dislocations
- ϕ is the angle between the interface and the normal to the slip plane

The results of Figure 7.13 show the expected resolution in x-ray peak position measurements as a function of the effective counts (a), FWHM (b), and Bragg angle (c) for epitaxial layers with Gaussian diffraction profiles. The results indicate that the x-ray resolution from peak position measurements is directly proportional to the FWHM and $\cot(\theta_B)$, but decreases monotonically with the number of effective counts; in other words, increasing the effective counts and Bragg angle improves the resolution, whereas an increase in the FWHM degrades the resolution. For the values of effective counts, FWHM, and Bragg angles explored, possible resolutions of 10^{-2} to 10^{-7} are attainable with x-ray measurements of peak position in epitaxial layers that exhibit a Gaussian diffraction profile.

Measurements of diffraction peak widths are commonly used to estimate dislocation densities in heteroepitaxial layers, and occasionally the variation of the peak broadening is used to detect the onset of lattice relaxation. It is therefore important to determine the resolution of peak width measurements. The results of Figure 7.14 show the dependence of the peak width resolution on the effective number of counts and peak FWHM with a Gaussian diffraction peak. Similar to the case of peak position measurements, the resolution is directly proportional to the FWHM and decreases monotonically with the effective number of counts. A wider selection of calculations for both types of measurements, and a treatment of both the Gaussian and Lorentzian cases, may be found in Appendix VIII.

7.2.6 Reciprocal Space Maps

In the mapping mode of operation for the triple-axis diffractometer, the specimen and analyzer axes are uncoupled. In this case, the specimen axis ω is scanned for each particular setting of ω′. This results in a three-dimensional map of the reflectivity versus ω and ω′. For any particular scan of ω with fixed ω′, the measured profile is given by

$$R_{\omega'}(\omega) = \int_{\alpha} R_m(\alpha) R_s(\omega - \alpha) R_a(2\omega - \omega') d\alpha. \tag{7.58}$$

Also, for a scan of ω′ with fixed ω, the measured profile is given by

$$R_{\omega}(\omega') = \int_{\omega} R_m(\alpha) R_s(\omega) R_a(2\omega - \omega') d\omega. \tag{7.59}$$

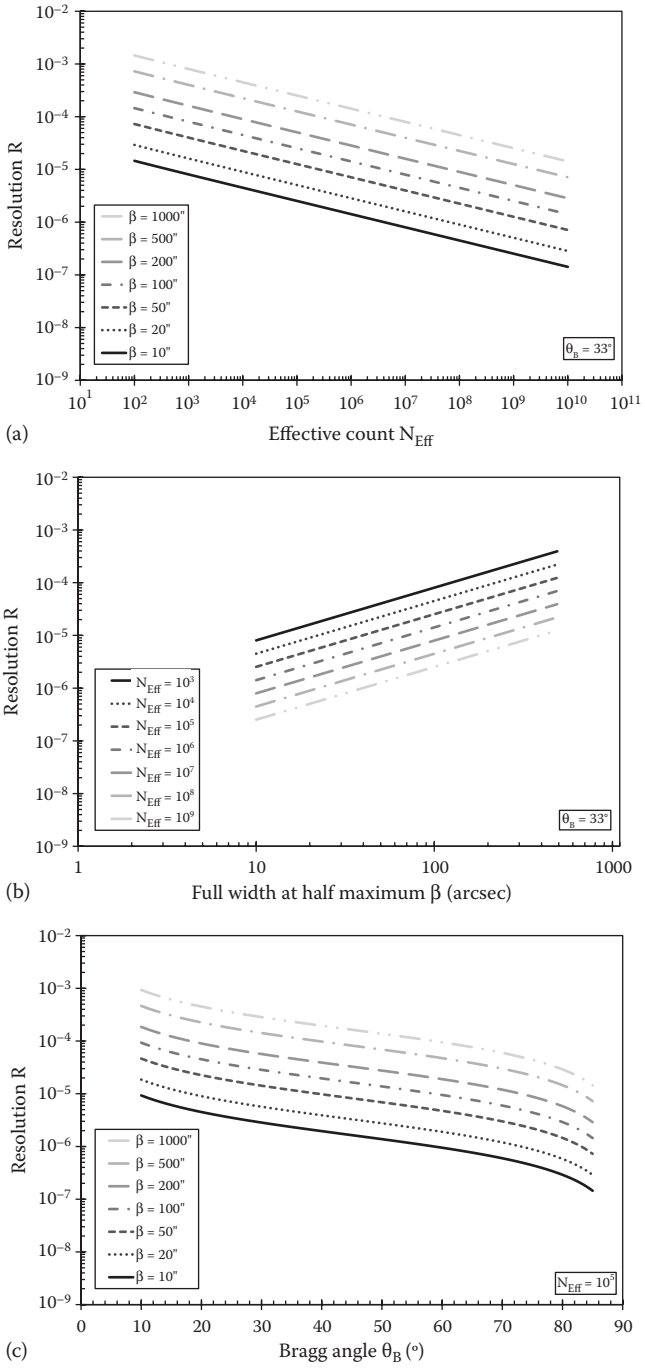


FIGURE 7.13
Resolution of XRD from measurements of a Gaussian peak position as a function of (a) the effective count with FWHM as a parameter and a Bragg angle of 33°, (b) the FWHM with the effective count as a parameter and a Bragg angle of 33°, and (c) the Bragg angle with FWHM as a parameter and an effective count of 105. The assumed Poisson ratio was ν = 1/3. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. Copyright 2015, World Scientific.)

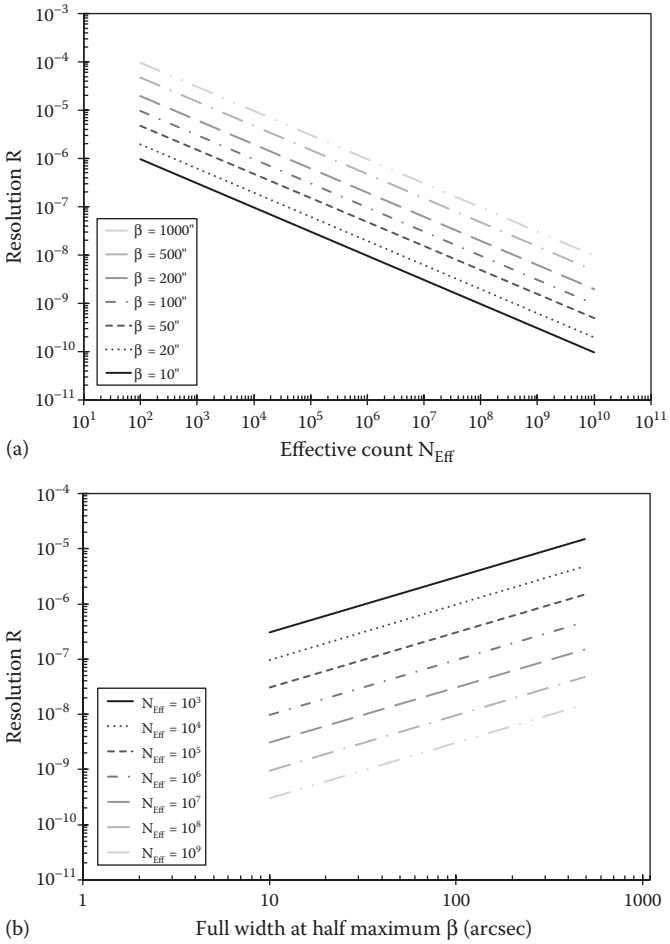


FIGURE 7.14
Resolution of XRD from measurements of a Gaussian peak width as a function of the (a) effective count for various FWHMs and (b) FWHM for various effective counts. The assumed number of threading dislocations is $x = 2$, and the angles associated with the slip plane are $\alpha,\phi = 60°$. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. Copyright 2015, World Scientific.)

The three-dimensional map obtained by varying both ω and ω′ contains the ω − 2ω scan as one cross section, and in practice, the mapping mode is often used to obtain this profile, in order to eliminate difficulties associated with the critical alignment of the analyzer crystal in the diffraction profile mode.

Commonly, reciprocal space maps (RSMs) are measured in the vicinity of the *hkl* reciprocal lattice point for one symmetric reflection and one or more asymmetric reflections, sometimes at different azimuths to aid in analyzing the crystallographic tilt. For cubic semiconductors grown with (001) orientation, the most commonly used symmetric reflection is 004, while 115, 224, 335, and 117 are popular asymmetric reflections. For nitride semiconductors grown with (0001) orientation, the 0002 symmetric reflection and the 10–13 or 1–214 asymmetric reflections may be employed.

In the presentation of the mapping results, it is common practice to render a contour map in two dimensions, and to translate the triple-axis diffraction map from the angular

coordinates ω′ and ω to the reciprocal space coordinates $\Delta q_{\|}$ and $\Delta q_{\perp}$, with units of nm^{-1}. Here, $\Delta q_{\|}$ and $\Delta q_{\perp}$ are the departures from the reciprocal space point in the in-plane and out-of-plane directions, respectively. The resulting map is called a "reciprocal space map."

The transformations from angle-scale coordinates to reciprocal space coordinates are given by[41]

$$\Delta q_{\perp} = \frac{4\pi}{\lambda}\left[n\delta(\omega - 2\omega')\cos\theta_B\right] \tag{7.60}$$

and

$$\Delta q_{\|} = \frac{4\pi}{\lambda}\left[m\delta\omega\sin\theta_B\right] \tag{7.61}$$

for small variations $\delta(\omega - 2\omega')$ and $\delta\omega$, where λ is the x-ray wavelength and θ_B is the Bragg angle, employed to measure the RSM. n and m are integers that represent the number of steps from the reciprocal lattice point in the *hkl* direction and the orthogonal direction, respectively, and the angular step sizes are $\delta(\omega - 2\omega')$ and $\delta\omega$, respectively.

It is unfortunate that no standard way of presenting RSM data has emerged. Sometimes, absolute scales are used for $q_{\|}$ and $q_{\perp}$, rather than relative scales for $\Delta q_{\|}$ and $\Delta q_{\perp}$. However, this is simply a change of scales on the RSM axes and the qualitative interpretation of the map is unchanged.

Reciprocal space mapping is a useful tool for the characterization of heteroepitaxial layers, because it allows the separation of strain broadening and angular broadening of defects, and lends itself readily to the extraction of tilts in individual layers. There is a wealth of information contained in the RSM, and its interpretation is sometimes complicated by the fact that no standard means for its presentation has emerged. Nonetheless, Section 7.8.4 will provide some examples of RSMs and discuss their interpretation.

7.3 Electron Diffraction

Electron diffraction techniques, especially reflection high-energy electron diffraction (RHEED) and low-energy electron diffraction (LEED), are important for the characterization of semiconductor surfaces. RHEED, in fact, is a critical *in situ* diagnostic tool for MBE growth. It allows the verification of a smooth, contaminant-free surface prior to growth, as is necessary for the epitaxy of high-quality material. It can also be used to determine the growth rate, composition, and growth mode *in situ*. Surface structure can be studied by RHEED as well, although LEED is used more commonly for this purpose. Electron diffraction techniques require a high vacuum, and therefore cannot be used for *in situ* diagnostics during vapor phase epitaxy.

7.3.1 Reflection High-Energy Electron Diffraction

In a typical RHEED experiment, a high-energy (10–100 keV) beam of electrons is incident on the sample surface at a shallow angle of 1°–2°. Diffraction of the electrons is governed by the Bragg law, as with XRD. However, there are two important differences between RHEED and

the x-ray case. First, the electrons do not penetrate significantly into the sample, so diffraction is essentially from the two-dimensional lattice on the surface. Second, for the high-energy electrons used in RHEED, the Ewald sphere is large in diameter, so many reflections are excited at once.

Because the diffraction occurs from a two-dimensional net of atoms on the surface, the reciprocal lattice comprises a set of rods perpendicular to the surface in real space. These rods can be indexed using the two Miller indices *hk*.

The electrons in a RHEED experiment behave as waves, with a de Broglie wavelength given by

$$\lambda = \frac{hc}{E}, \tag{7.62}$$

where:

- h is the Planck constant
- c is the speed of light
- E is the electron energy

For example, an electron energy of 100 keV corresponds to a de Broglie wavelength of 3.7 pm. The radius of the Ewald sphere is $|k_0| = 1700\ \text{nm}^{-1}$, whereas the separation of the rods in the reciprocal lattice $2\pi/a$ might be about $20\ \text{nm}^{-1}$. The Ewald sphere is so large compared with the separation of the reciprocal lattice rods that it will intersect several rods, exciting several Bragg reflections for any given geometry. The diffraction pattern therefore comprises a set of streaks, as shown in Figure 7.15.

If the distance from the specimen to the screen is L and the separation between the 00 and *hk* streaks on the screen is t, then

$$t = L\tan\left(2\theta_{hk}\right). \tag{7.63}$$

If the surface structure is a square lattice with lattice constant a, then by the Bragg law,

$$\lambda = 2d\sin\theta_{hk} = \frac{2a\sin\theta_{hk}}{\left(h^2 + k^2\right)^{1/2}}. \tag{7.64}$$

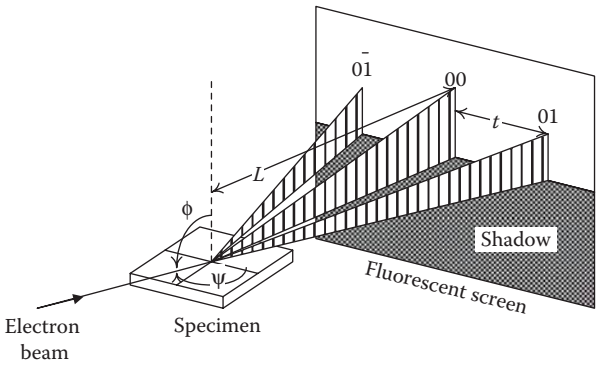


FIGURE 7.15
RHEED experiment.

In a RHEED experiment, $\lambda \ll a$, so that

$$a \approx \left(h^2 + k^2\right)^{1/2} \lambda L / t. \tag{7.65}$$

Therefore, the lattice constant for the surface may be determined. This analysis can be extended to other surface lattices, and by performing RHEED experiments at different azimuths ψ, it is possible to determine the dimensions of the surface unit mesh. In principle, it is also possible to determine the surface structure (the positions of the atoms in the unit mesh), although this is not usually done.

RHEED is commonly used *in situ* during MBE growth to discern the growth rate and growth mode. The growth rate may be determined from RHEED intensity oscillations, for which the period corresponds to one monolayer of growth. The surface roughness, and therefore the growth mode, may be discerned from the nature of the RHEED pattern. As noted previously, a streaky pattern is indication of an atomically flat surface. In the case of a rough surface, the electron beam will penetrate islands or other structures on the surface, giving rise to diffraction from a three-dimensional lattice. Therefore, the RHEED pattern becomes spotty in this case.

7.3.2 Low-Energy Electron Diffraction

LEED is an electron diffraction method that utilizes a beam of low-energy (<1 keV) electrons at normal incidence, as shown in Figure 7.16. Here, too, it can be assumed that only the top layer of atoms gives rise to the diffraction. Due to the normal incidence, however, the reciprocal lattice rods are nearly perpendicular to the (approximately flat) Ewald sphere surface. This gives rise to a pattern of spots on the pattern corresponding to the intersection of the reciprocal lattice rods with this Ewald sphere. The positions of these spots can be used to determine the unit cell of the surface mesh, and the simulation of the spot intensities makes it possible to determine the positions of the atoms within the unit cell. LEED is therefore used to study surface structure and surface reconstructions.

Figure 7.17 shows example LEED patterns for a Si(111) surface that has been exposed to In, inducing a Si(111) (4 × 1)–In structure.[46] The left pattern was obtained using an electron energy of 70 eV, and the 4 × 1 unit cell has been highlighted. The pattern on the right was

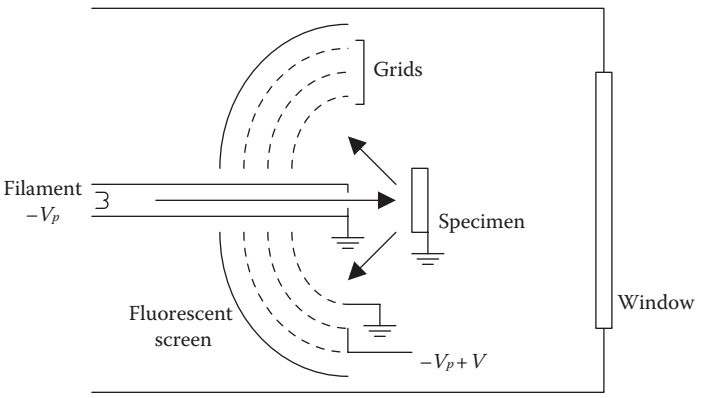


FIGURE 7.16
LEED experiment.

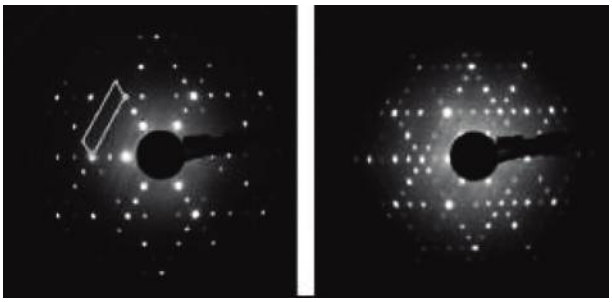


FIGURE 7.17
LEED patterns for the Si(111) (4 × 1)–In surface. The left pattern was obtained using an electron energy of 70 eV, and the 4 × 1 unit cell has been highlighted. The pattern on the right was obtained using an electron energy of 120 eV. With the higher electron energy, the diffraction spots are more closely spaced. (Reprinted from J. Wang et al., *Phys. Rev. B*, 72, 245324 [2005]. With permission. Copyright 2005, American Physical Society.)

obtained using an electron energy of 120 eV. With the higher electron energy, the diffraction spots are more closely spaced.

7.4 Microscopy

Microscopic methods include optical microscopy (OM), scanning electron microscopy (SEM), transmission electron microscopy (TEM), atomic force microscopy (AFM), and scanning tunneling electron microscopy (STEM). Most of these are used to image the surface, and thus characterize the surface morphology and growth mode. They may also be used in conjunction with crystallographic etches to determine defect densities. AFM has sufficient resolution for the study of the surface atomic structure. TEM enables the imaging of dislocations and other crystal defects within the volume of the crystal; it is therefore important in the study of dislocations and lattice relaxation in mismatched heteroepitaxial layers.

7.4.1 Optical Microscopy

OM is routinely used to characterize the surface morphology of heteroepitaxial layers, because it is rapid and nondestructive. However, the method offers only modest resolution and depth of field. The lateral resolution r_{lateral} may be estimated by the Rayleigh criterion to be

$$r_{\text{lateral}} = \frac{0.6\lambda}{NA}, \tag{7.66}$$

where:

λ is the optical wavelength
NA is the numerical aperture of the objective lens

It can be improved by increasing the numerical aperture. However, this involves a tradeoff with the depth of field, or axial resolution r_{axial}, which can be estimated as

$$r_{\text{axial}} = \frac{1.4\lambda n}{NA^2}, \tag{7.67}$$

where n is the index of refraction. Typical objectives have numerical apertures of 0.1–1, so both the lateral resolution and depth of field are measured in microns.

Nomarski interference contrast microscopy[47] is often used for the microscopic inspection of heteroepitaxial layers, for the evaluation of the surface morphology, or for the counting of etch pits after the use of a crystallographic etch. This method produces an image from the gradient of the refractive index, and therefore acts as a high-pass filter that accentuates edges and boundaries.

7.4.2 Transmission Electron Microscopy

TEM[48] is a valuable technique for the observation of dislocations, stacking faults, twin boundaries, and other crystal defects in heteroepitaxial layers. TEM characterization is applicable to most heteroepitaxial semiconductor samples, provided that they can be thinned to transmit electrons, and that they are stable when exposed to a high-energy electron beam in an ultrahigh vacuum. Conventional TEMs use electron energies of ~100 keV, whereas this number can be greater than 1 MeV in a high-voltage TEM. For observation in a conventional TEM, typical heteroepitaxial semiconductor samples must be thinned to less than about 100 nm. This requirement may be relaxed somewhat if the sample is made up of light atoms with low atomic number (such as Si, SiC, or sapphire), or if high-voltage electrons are used. The sample preparation is destructive, and in some cases may alter the defects that are to be observed.

The electrons in a TEM behave as waves, with a de Broglie wavelength given by

$$\lambda = \frac{hc}{E}, \tag{7.68}$$

where:

- h is the Planck constant
- c is the speed of light
- E is the electron energy

For example, an electron energy of 100 keV corresponds to a de Broglie wavelength of 3.7 pm. The TEM uses lenses to produce an image, but the lenses are electromagnetic in nature. Lens aberrations, along with mechanical and electrical instabilities, usually limit the resolution of the TEM to 2 Å.

The operation of a TEM instrument is shown schematically in Figure 7.18. Collimated high-energy electrons from a condenser lens impinge on the semiconductor specimen and are transmitted through it. The electrons are scattered into particular directions by the crystalline sample according to the Bragg law for diffraction. These diffracted beams are brought into focus at the focal plane for the objective lens.

In the diffraction mode, the first intermediate lens is focused on the back focal plane of the objective lens, thus capturing the diffraction pattern. This diffraction pattern is magnified and projected by the combination of the intermediate and projection lenses. The diffraction pattern displayed on the screen comprises an array of spots, each corresponding to a particular diffraction vector **g**. The diffraction mode is used to index the diffraction

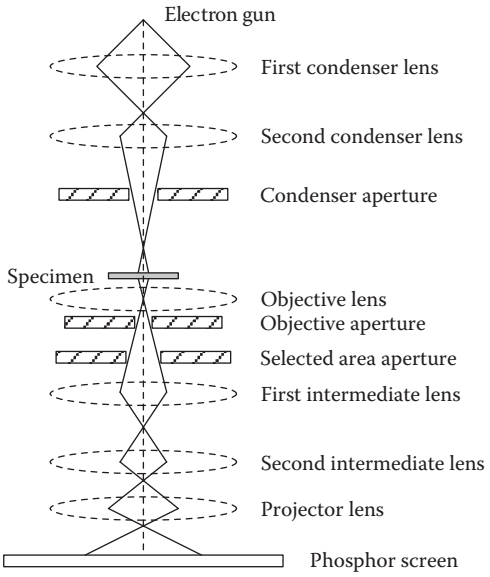


FIGURE 7.18
Transmission electron microscope.

beams and to facilitate the selection of the diffraction spot(s) to be used in ultimately forming an image.

In the imaging mode, the intermediate lens is focused on the inverted image of the sample formed by the objective lens. This image is magnified and projected onto the screen with an overall magnification of up to 10^6. An aperture at the back focal plane of the objective lens is used to select only one diffracted beam to form the image. If the beam transmitted directly through the image $\mathbf{g} = [000]$ is chosen, a "bright-field" image results. If one of the diffracted beams is chosen to form the image, then a "dark-field" image is produced.

The variation of the image intensity leaving the specimen may be understood using two simplifying approximations. First, the specimen is assumed to behave as if made up of narrow columns (the "column approximation") with axes parallel to the incident beam. The image therefore represents an intensity bit map for the array of columns. Second, it is assumed that the image is formed by the directly transmitted beam plus only one diffracted beam (the "two-beam approximation").

A uniform, perfect crystal, with uniform thickness, will produce an image with uniform electron intensity. Image contrast results from crystal nonuniformities, including variations in thickness, changes in composition, inclusions, and voids. Dislocations may also produce image contrast, if they displace the diffracting planes such that their separation or orientation changes. Based on the column approximation, we can state that the condition for image contrast by a crystal defect is $\mathbf{g} \cdot \mathbf{u} \neq 0$, where $\mathbf{u}$ is the vector by which atoms are displaced from their normal sites within a particular column. Put another way, the condition for invisibility is $\mathbf{g} \cdot \mathbf{u} = 0$.

For an edge or screw dislocation, image contrast will result if the Burgers vector has a component in the direction of the diffraction vector. In other words, the condition for invisibility (zero contrast) for an edge or screw dislocation is

$$\mathbf{g} \cdot \mathbf{b} = 0 \text{ (invisibility condition).} \tag{7.69}$$

For dislocations of mixed character, there is no condition for which $\mathbf{g} \cdot \mathbf{u} = 0$. Instead, the invisibility criterion is *approximately satisfied* for $\mathbf{g} \cdot \mathbf{b} = 0$ and weak contrast is observed even if $\mathbf{g} \cdot \mathbf{b} = 0$.

The invisibility condition can be applied to the determination of the Burgers vector direction for a dislocation. If a dislocation is invisible (or nearly invisible, in the case of a mixed dislocation) in two images produced using the diffraction vectors $\mathbf{g}_1$ and $\mathbf{g}_2$, then its Burgers vector must be perpendicular to both diffraction vectors. This means that its Burgers vector is in the direction $\mathbf{g}_1 \times \mathbf{g}_2$. In this way, the Burgers vector and the character (edge, screw, or mixed) may be determined for dislocations in a heteroepitaxial layer.

A critical step in any TEM experiment is the sample preparation. It is necessary to prepare a thin foil that includes the region to be examined (e.g., the heterointerface) and which is thin enough to transmit the electrons. Often, this can be achieved by a combination of wet etching and ion milling. In some cases, etch stop layers can be used for the preparation of thin foils. For example, plan-view samples of heteroepitaxial zinc blende layers on GaAs can be prepared by the use of a thin etch release layer such as AlAs. Selective etching of the AlAs layer allows release of the heteroepitaxial layer(s) above, which can be floated on water and picked up by a carbon-coated TEM grid for microscopic investigation.[49]

If the results from TEM examination are to be meaningful, the sample being studied must be stable under irradiation by a high-energy electron beam. This condition is not always met in heteroepitaxial samples. For example, in the case of $In_xGa_{1-x}As$, it has been found that electron irradiation of the sample can excite motion of glissile dislocations, and that the glide can be started or stopped by condensing or expanding the electron beam.[49]

7.4.3 Scanning Tunneling Microscopy

The scanning tunneling microscope is an instrument that can measure surfaces with atomic-scale resolution. It was invented by Binnig and Rohrer in 1981, for which they received the Nobel Prize in 1986. The basic principle of operation can be understood using Figure 7.19. Here, P_x and P_y are piezoelectric elements that allow a metal tip to be scanned over a surface with Angstrom positioning accuracy. A feedback control unit (CU) biases the

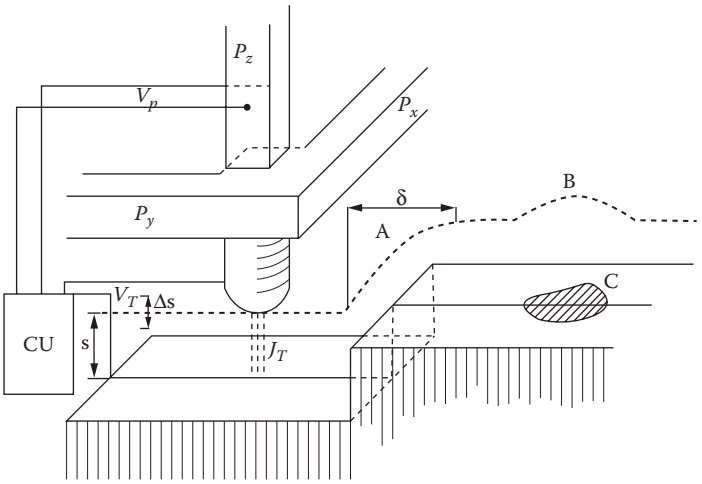


FIGURE 7.19
Scanning tunneling microscope. (Reprinted from G. Binnig et al., *Phys. Rev. Lett.*, 49, 57 [1982]. With permission. Copyright 1982, American Physical Society.)

third piezoelectric element such that a constant tunneling current flows between the tip and the surface. Because the tunneling current varies exponentially with the tip-to-surface separation, the tip will follow the contour of the surface during the scan. The deflections of the three piezoelectric elements are proportional to the three bias voltages V_z, V_x, and V_y. Therefore, a map of V_z as a function of V_x and V_y gives a topographic map of the surface.

The tunnel current density flowing from the tip to the surface depends strongly on the work function of the surface ψ and the tip-to-surface separation s and is proportional to

$$J_T \propto \exp\left(-A\psi^{1/2}s\right), \tag{7.70}$$

where $A = 1.025$ Å^{-1} eV$^{-1/2}$. Therefore, for a typical surface work function of ~1 eV, a single atomic step on the surface would change the tunneling current by three orders of magnitude, in the absence of vertical adjustment by P_z. This leads to a vertical resolution of 0.2 Å.

As shown in Figure 7.19, false surface features can emerge as a consequence of surface contaminants that modify the work function (C in Figure 7.19). However, the work function mimicked features can be separated from true surface structures by modulating the tip distance during the scan.

Even though practical probe tips will tend to be blunt and irregular in shape (see Figure 7.20), the exponential dependence of the tunneling current on tip separation causes a localized fine tip, or even a single atom on it, to be active in the tunneling. Therefore, the lateral resolution of 10 Å can be achieved.

The STM provides atomic-scale images of surfaces, and can be used to study surface steps and kinks, the growth mode, and the earliest stages of islanding. Also, because there are surface trenches associated with subsurface misfit dislocations in very thin layers, STM can be used to study the initiation of lattice relaxation in highly mismatched heteroepitaxial layers. A drawback of this instrument, however, is the necessity for a conducting specimen. The surfaces of insulators may not be investigated due to charging effects. Moreover, semiconductors with high resistivity or native oxide films cannot be examined by STM.

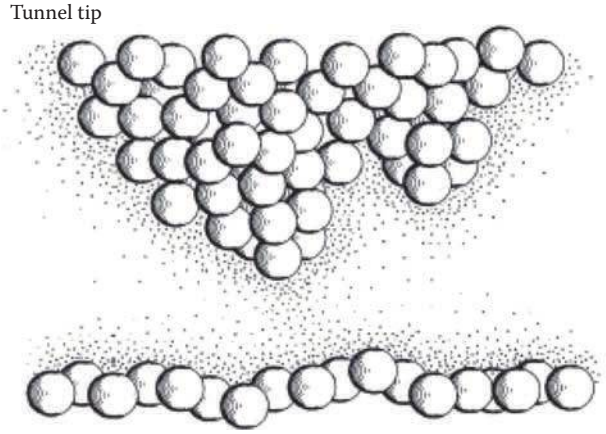


FIGURE 7.20
STM tip over a surface. The tip may be blunt and irregular; however, the exponential dependence of the tunneling current causes one microtip to be active. In this picture, the upper microtip is expected to give rise to a tunneling current 1/1000th of the tunneling current associated with the lower tip. (Reprinted from G. Binnig and H. Rohrer, *Rev. Mod. Phys.*, 59, 615 [1987]. With permission. Copyright 1987, American Physical Society.)

7.4.4 Atomic Force Microscopy

The atomic force microscope is an instrument with similar resolution and applications as the STM, but it can be used with insulating surfaces. The AFM, first proposed by Binnig et al.[50] in 1986, combines the principles of the STM and the stylus profilometer, resulting in an atomic-scale profilometer.

The operation of the AFM can be understood with the aid of Figure 7.21. A cantilever with a sharp tip (which need not be conducting) is placed between the STM tunneling tip and the sample to be examined by AFM. While scanning, a very small and constant force is maintained on the AFM tip. This can be done by different means, in one of several operating modes. In one mode of operation, the force exerted on the AFM stylus by its piezoelectric element is adjusted for constant tunneling current in the STM. Regardless of the details of how the feedback system is employed, the z displacement of the STM corresponds to the z displacement of the surface being examined by the AFM.

Some of the capabilities of this characterization technique are apparent in the AFM micrographs of Figure 7.22. The images show surfaces of $In_{0.65}Ga_{0.35}As$ layers of different thicknesses, grown on InP(001) substrates by metalorganic vapor phase epitaxy (MOVPE) and reported by Jasik et al.[51] The layer of Figure 7.22a is 23.4 Å thick. Monolayer surface steps can be seen clearly, with an average step spacing of 200 nm, corresponding to an offcut angle of about 0.25°. The layer of Figure 7.22b is 103.4 Å thick and displays a more irregular arrangement of surface steps. The new features that have appeared on the 125 Å thick layer of Figure 7.22c are associated with misfit dislocations at the interface.

7.5 Crystallographic Etching Techniques

Crystallographic etches can be used to reveal crystal defects, such as dislocations or stacking faults. They are also useful for the delineation of interfaces and p-n junctions. Moreover, crystallographic etches produce pits with orientation-dependent characteristics, so they can be used to identify crystal directions and detect inversion domains. In all of these applications, the usefulness of the etch comes about because of its sensitivity to the surface; defects, interfaces, or junctions are revealed due to the modified etch rate in their vicinity.

Quite generally, an etching process may be polishing or crystallographic. Either type of process involves the transport of reactants through a diffusion boundary layer, a reaction or reactions at the surface, and transport of the reaction products away from the etching surface. If the etch rate is limited by the diffusion of either reactants to the surface or the products away from the surface, it is termed *diffusion limited* or *mass transport controlled*. A diffusion-limited etch is insensitive to the crystallographic orientation and also local variations in the surface structure; it is therefore polishing. On the other hand, a reaction-rate-limited etch is sensitive to the surface, and will be a crystallographic etch. The etch rate in such a case will vary across the surface, and can be used to reveal defects such as dislocations.

Wet chemistry is used almost exclusively for crystallographic etching. However, gaseous etching processes may also be reaction rate limited under the proper conditions of temperature and flow rates. For example, Tachikawa and Mori[52] demonstrated the use of HCl-GaCl for crystallographic etching of GaAs/Si(001) and GaP/Si(001) in the growth chamber,

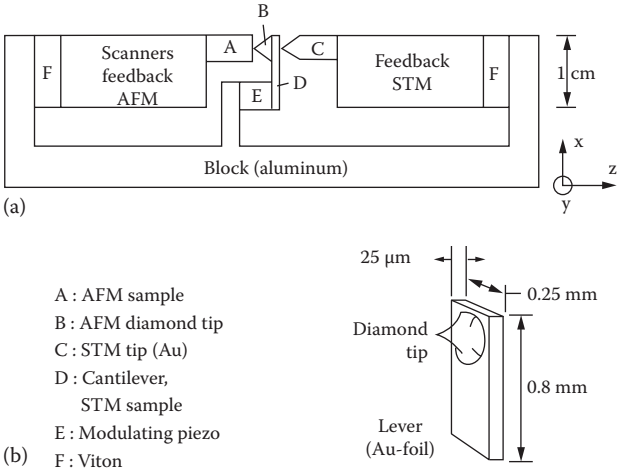


FIGURE 7.21
Atomic force microscope. (a) System schematic. (The tip is not shown to scale.) (b) Cantilever with diamond tip, showing the dimensions. (Reprinted from G. Binnig et al., *Phys. Rev. Lett.*, 56, 930 [1986]. With permission. Copyright 1986, American Physical Society.)

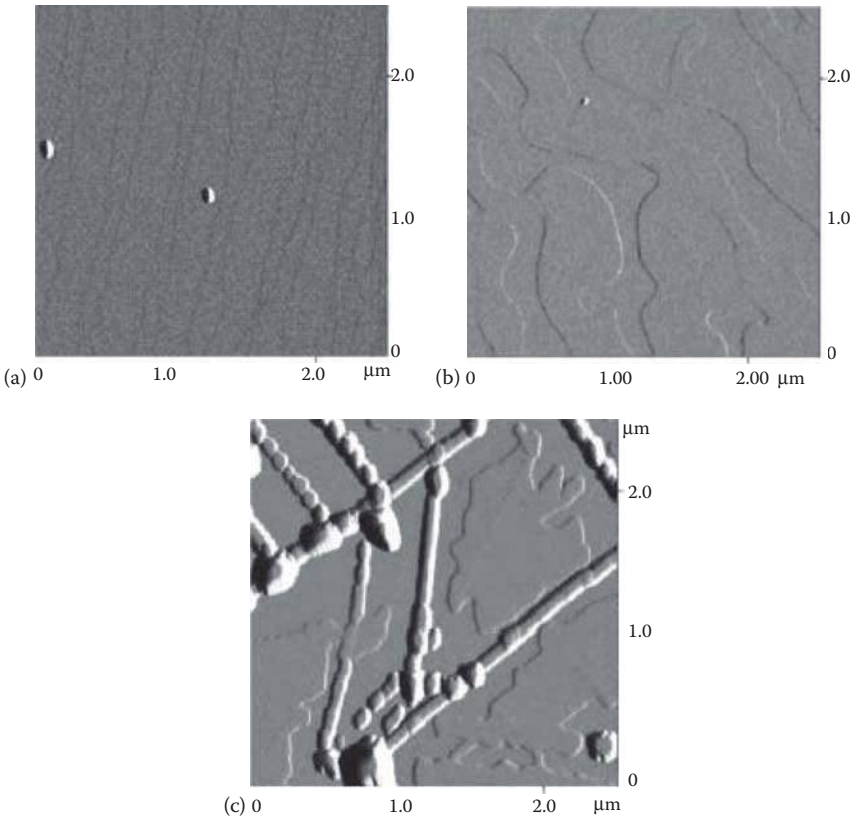


FIGURE 7.22
Surface steps as observed by AFM on a 23.4 Å thick layer of $In_{0.65}Ga_{0.35}As$ on InP(001). (Reprinted from A. Jasik et al., *Thin Solid Films*, 412, 50 [2002]. With permission. Copyright 2002, Elsevier.)

at the growth temperature. As another example, Ohno et al.[53] found that a hydrogen plasma revealed dislocation etch pits on the surface of ZnSe.

Wet chemical etches usually comprise an oxidizer, a complexing agent, and a diluent. The most commonly used oxidizers are H_2O_2 and HNO_3. The complexer reacts with the oxidized surface to create a water-soluble complex. While HF is the most common complexer, nitric, sulfuric, phosphoric, and citric acid are also sometimes used. Br_2, which is used in some etch formulations, serves the dual role of oxidizer and complexer. The diluent, usually water or CH_3COOH, is sometimes omitted.

The most common application for crystallographic etching is the evaluation of dislocation densities in bulk or heteroepitaxial semiconductors. Here, the points where dislocations emerge at the surface are marked by the appearance of hillocks or, more commonly, pits. These features occur due to the reduced or enhanced etch rate in the strained region around the dislocation.

Etch pits on the surface of a semiconductor crystal usually reveal the crystal symmetry and can be used to determine the orientation. For example, Si(111) treated by Sirtl etch shows triangular pits, while molten KOH etching of 6H-SiC(0001) reveals hexagonal pits. On the other hand, molten KOH etching of GaAs(001) produces approximately rectangular pits, which are elongated along the [110] direction. This makes it possible to distinguish the [110] and [$\bar{1}$10] directions in the surface, and therefore find whether inversion domains are present.

Crystallographic etching has also been used extensively to delineate p-n junctions or other interfaces in multilayered heteroepitaxial structures. For example, dilute (15%) A-B etch in water etches p-type material much faster than n-type material and has been used to delineate p-type regions in cleaved AlGaAs/GaAs laser structures.[54] A modified A-B etch has also been used to delineate compositional steps on the cleaved edge of a GaAsP/GaAs/Ge(001) heterostructure.[55]

The application of crystallographic etching to the determination of threading dislocation densities is detailed in Section 7.11.2. The compositions of some commonly used crystallographic etches are tabulated in Appendix VI.

7.6 Photoluminescence

Photoluminescence (PL) is commonly employed *ex situ* to assess the suitability of heteroepitaxial structures for optoelectronic devices such as light-emitting diodes (LEDs) and laser diodes. A wealth of information may be obtained from PL spectra, especially by taking measurements at different temperatures or with different excitation wavelengths or intensities. Much of this information is particularly useful in studies of doping, which are beyond the scope of this book and will not be elaborated here. It is also possible to use PL for the determination of structural information, such as compositions and strains in heteroepitaxial layers. The application to quaternary layers is of special interest; here, PL is commonly used in conjunction with HRXRD for determination of the composition and strain. In a nonalloyed semiconductor such as GaAs, the strain may be determined directly from PL measurements by using the strain-induced splitting of the light hole and heavy hole bands, and this technique has been applied to GaAs on Si.[56] Photoluminescence microscopy (PLM) is an imaging technique that creates a map of the PL intensity. This can be used to study electronically active defects such as dislocations, which show up as dark regions on PLM images.

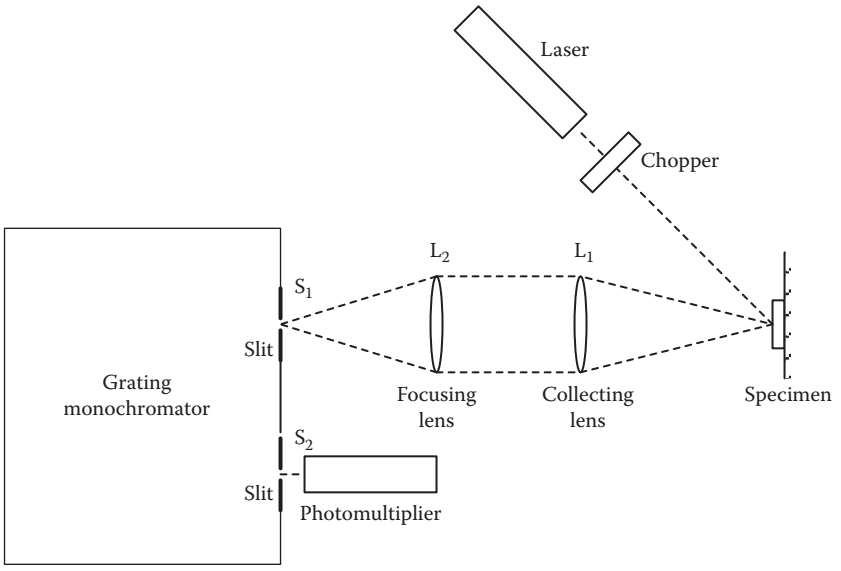


FIGURE 7.23
Photoluminescence setup.

A typical PL setup is shown schematically in Figure 7.23. A laser with above-bandgap photons excites electron–hole pairs in the sample. Recombination of the excess carriers gives rise to the emission of characteristic wavelengths associated with the electronic transitions in the sample. The emitted radiation (which may span the range from ultraviolet to infrared, depending on the specimen) is collimated by a collecting lens L_1 and then focused on the entrance slit of the monochromator by a second lens L_2. The monochromator is scanned in wavelength during the experiment, so an intensity versus wavelength spectrum is obtained. At each wavelength, the intensity passing through the monochromator is measured using a photomultiplier tube. Usually, the exciting laser beam is chopped to facilitate phase-sensitive detection with a lock-in amplifier; this greatly improves the signal-to-noise ratio.

Luminescence imaging techniques are important for the study of defects in heteroepitaxial layers. This is because enhanced nonradiative recombination of carriers occurs in the vicinity of electronically active defects such as dislocations. They therefore show up as darkened regions on a map of luminescence intensity.

PLM is an imaging technique that can be performed in a "micro-PL" system or a scanning near-field optical microscope (SNOM). As with standard PL, the PLM technique involves the excitation of electron–hole pairs in the sample by laser illumination, the recombination of the resulting free carriers, and the detection of emitted light. However, PLM requires spatial resolution of the signal so that an image can be obtained. Typically, a microimaging system with a resolution of 1 μm is used.[57]

In a micro-PL system, the laser excitation is spot focused (typical spot size of ~50 μm) to allow illumination of a small area of the sample. In the case of multilayer structures, the incident wavelength can be chosen to stimulate PL from only the layer(s) with the smallest bandgaps. The resulting images are captured using a digital camera. Typically, cryogenic temperatures (e.g., 77 K) are used with the advantage of one to three orders of magnitude increase in the PL intensity.

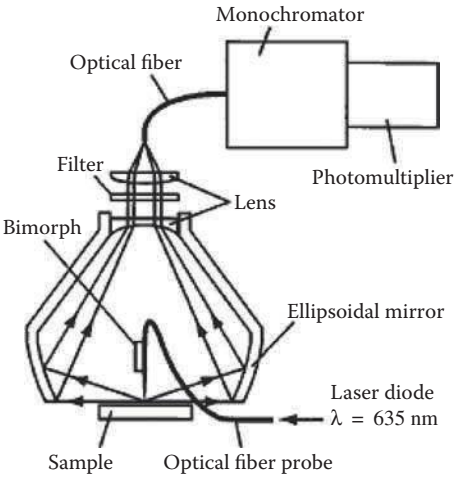


FIGURE 7.24
PL microscopy setup. (Reprinted from Y. Ohizumi et al., *J. Appl. Phys.*, 92, 2385 [2002]. With permission. Copyright 2002, American Institute of Physics.)

PLM images may also be obtained using a SNOM.[58] In this case, the laser illumination and the collected light both pass through optical fibers. Using specially prepared metal-coated fibers with tapered tips,[59] images with submicron resolution can be obtained. A typical SNOM-based PLM arrangement is shown in Figure 7.24. The excitation comes from a laser diode coupled with an optical fiber. This can be scanned over the surface of the specimen. The luminescence is collected by an ellipsoidal mirror and a lens, and then fed to the monochromator through a second optical fiber. Scanning of the excitation source fiber aperture in x and y allows the creation of a PL intensity map. The monochromator is fixed upon a particular wavelength for the measurement of the PLM image. Often, this is the wavelength corresponding to peak intensity in the PL spectrum. Sometimes, other wavelengths are chosen in order to study subtle aspects of the electronic behavior of defects.

The depth sensitivity of PLM may be tailored by adjustment of the excitation wavelength.[60] Generally, a shorter wavelength will have a larger absorption coefficient and will be absorbed close to the sample surface. Therefore, the resulting PLM image will be associated with photons emitted from the material near the surface. On the other hand, a longer wavelength will penetrate more deeply into the sample.

Figure 7.25 shows the PLM image for a GaAs/$In_{0.15}Ga_{0.85}As$ multiquantum well structure (seven periods, 50 nm GaAs and 16 nm $In_{0.15}Ga_{0.85}As$) grown on a GaAs(001) substrate by MBE.[61] Dark lines corresponding to misfit dislocations are seen to run parallel to the [110] and [$\bar{1}$10] directions. This image was obtained using the peak emission wavelength from the PL spectrum, and a bandwidth of 60 meV. The areas of the dark lines exhibit a 25% reduction in PL intensity compared with the surrounding regions.

Cathodoluminescence is another imaging technique that can be used to study defects in heteroepitaxial structures. This method bears many similarities to PLM, except that it is conducted in a scanning electron microscope using an electron beam for the excitation.

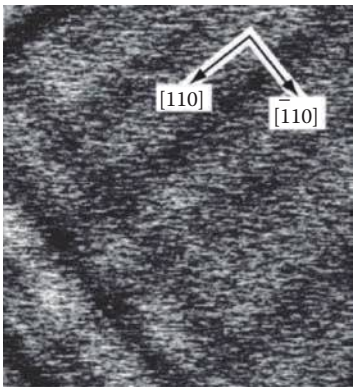


FIGURE 7.25
PLM image for a GaAs/In0.15Ga0.85As multiquantum well structure (seven periods, 50 nm GaAs and 16 nm In0.15Ga0.85As) grown on a GaAs(001) substrate by MBE. (Reprinted from Y. Ohizumi et al., *J. Appl. Phys.*, 92, 2385 [2002]. With permission. Copyright 2002, American Institute of Physics.)

7.7 Growth Rate and Layer Thickness

Layer thicknesses and growth rates have been determined by many means; some are simple (gravimetric measurements), whereas others are quite sophisticated (RHEED oscillations, Fourier transform infrared [FTIR] spectroscopy, x-ray Pendellosung).

RHEED provides a means for the *in situ* determination of growth rates and layer thicknesses in an MBE growth chamber. Because the electron beam is incident on the sample at a very shallow angle (typically 1°–2°), RHEED analysis can be performed while the heteroepitaxial layer is growing. In the case of Frank–van der Merwe (layer-by-layer) growth, a streaky diffraction pattern is obtained, and the intensity of a particular streak oscillates with time. Much information is contained in the RHEED oscillation characteristics, and sophisticated models have been developed for their analysis. To determine the growth rate and layer thickness, it suffices to recognize that one period of the RHEED oscillations corresponds to the growth of 1 monolayer (ML). The intensity is maximum for a smooth surface. The nucleation of a new layer on this surface causes its roughening until the new layer completes, and then the process repeats. If the period of the RHEED oscillations is T, then the growth rate is

$$g = \frac{1ML}{T}. \tag{7.71}$$

Figure 7.26 shows representative RHEED oscillations measured during the MBE growth of GaAs(001) with incident angles of 1.33° and 0.93°, as indicated.[62] Here, the period of oscillations is 2.2 s, corresponding to a growth rate of 0.45 ML s^{-1} or 230 nm h^{-1}. Usually, the RHEED oscillations decay after only a few monolayers; in Figure 7.26, only about eight periods of oscillation are observed.

A convenient method for rapid, nondestructive layer thickness measurement is based on the reflectance characteristic, measured with an FTIR spectrometer. The reflectance (or transmittance) curve will contain an interference pattern, due to the interference of the

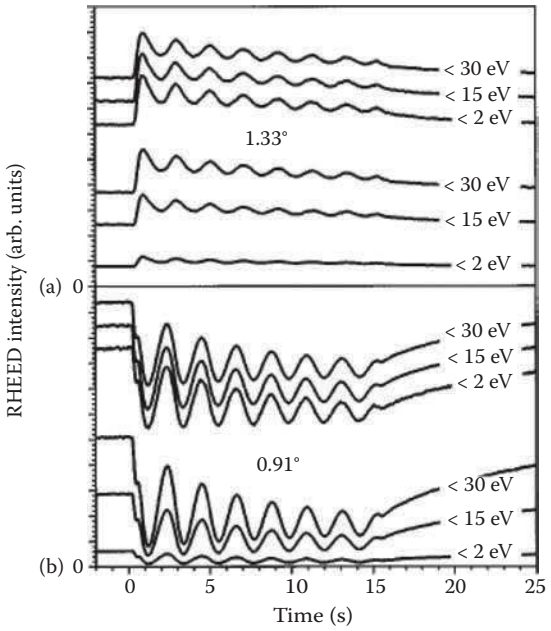


FIGURE 7.26
RHEED oscillations observed during the homoepitaxial growth of GaAs(001) on a 2 × 4 reconstructed surface. The primary beam energy was 20 keV. (a) Original curves and (b) normalized curves. Here, an energy filter was applied in front of the RHEED screen; the filter settings were as shown. (Reprinted from W. Braun et al., *J. Vac. Sci. Technol. B*, 16, 2404 [1998]. With permission. Copyright 1998, American Vacuum Society.)

waves reflected at the epitaxial layer surface and the layer–substrate interface, as long as the epitaxial layer and the substrate have different indices of refraction. In practice, this condition is satisfied even in the case of homoepitaxy due to the change in doping at the interface. If the measured interference fringe pattern contains m periods in the range of wavenumber from ν_1 to ν_2, the layer thickness is given by

$$h = \frac{m}{2\sqrt{n^2 - \sin^2\theta}} \left(\frac{1}{\frac{1}{\nu_1} - \frac{1}{\nu_2}} \right), \tag{7.72}$$

where:

n is the index of refraction for the heteroepitaxial layer
θ is the angle of incidence

Figure 7.27 shows example FTIR spectra for silicon epitaxial layers having different thicknesses.

For layers having high crystal perfection and smooth surfaces, the thickness can be determined from the x-ray rocking curve using the Pendellosung interference fringes. The spacing of the Pendellosung is given by

$$S = \frac{\lambda \sin\left(\theta_B \pm \phi\right)}{h \sin\left(2\theta_B\right)}, \tag{7.73}$$

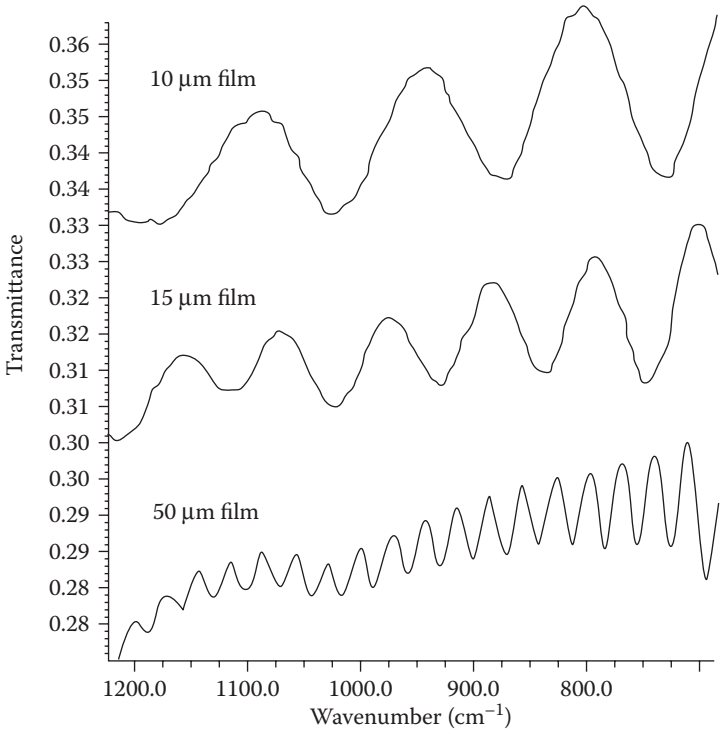


FIGURE 7.27
FTIR spectra measured for three Si epitaxial layers of different thicknesses, as indicated. (Courtesy of Thermo Electron Corporation, Madison, WI.)

where:

- λ is the x-ray wavelength
- θ_B is the Bragg angle
- ϕ is the angle between the diffracting planes and the surface ($\theta_B \pm \phi$ is the angle of incidence)
- h is the layer thickness

For example, for a symmetric 004 reflection from a (001) layer of GaAs, 1 μm thick, the Pendellosung spacing is expected to be 28 arcsec. Pendellosung can only be observed with a small-area x-ray beam. Otherwise, thickness nonuniformities or bending will suppress the fringing. Usually, Pendellosung are observed only for thin layers, whereas in thicker layers the nonuniformities, surface roughening, or defects will extinguish them. Figure 7.28 shows a simulated rocking curve for 0.2 μm AlAs on GaAs(001), in which the Pendellosung are clearly seen.

Layer thicknesses can be determined directly using TEM cross-sectional micrographs. In a multilayer structure, there is usually sufficient contrast so that all individual layers can be distinguished. In partially relaxed structures, defects enhance the visibility of interfaces. Marker layers (e.g., AlAs in GaAs, Ge in SiGe) can be inserted to delineate a particular point in the growth process, which is a unique feature associated with TEM.

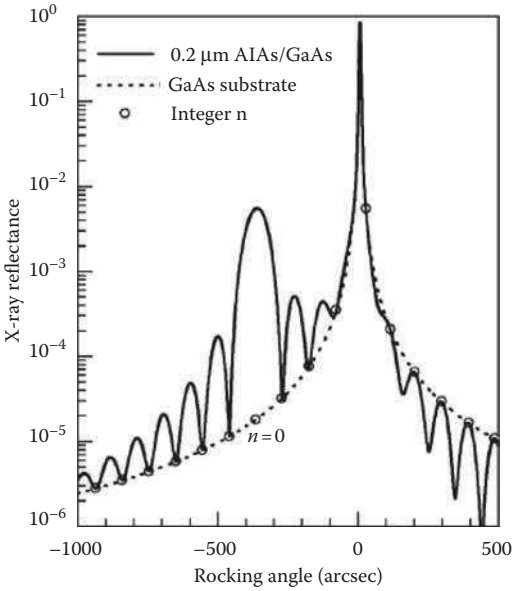


FIGURE 7.28
Simulated x-ray rocking curve for 0.2 μm AlAs on GaAs(001) using Cu kα_1 radiation. The spacing of the Pendellosung interference fringes is inversely proportional to the epitaxial layer thickness. (Reprinted from I. Kim et al., *J. Appl. Phys.*, 83, 3932 [1998]. With permission. Copyright 1998, American Institute of Physics.)

7.8 Determination of Composition and Strain

The strain in a single heteroepitaxial layer is most easily determined using double-axis XRD (double-crystal or Bartels diffractometer). It is usually appropriate to assume the strain is constant with depth. Then, for a binary layer, the relaxed lattice constant is known, so there is only one independent unknown. (Once the in-plane or out-of plane lattice constant is known, the other may be calculated.) In a strained ternary alloy (or the alloy $Si_{1-x}Ge_x$), there are two independent unknowns. One possible set is the relaxed lattice constant and the in-plane lattice constant. Thus, two measurements are necessary to characterize the layer. The number of required measurements increases with the number of degrees of freedom in the layer. Also, if there is a crystallographic tilt between the epitaxial layer and the substrate, this introduces two more unknown variables (the magnitude and direction of tilt), thus necessitating additional measurements. This section details the application of x-ray measurements to characterize the composition and strain in these heteroepitaxial layers.

7.8.1 X-Ray Diffraction Analysis of a Binary Heteroepitaxial Layer

In a nonalloyed heteroepitaxial layer such as Ge or GaAs (referred to as a binary layer here), the relaxed lattice constant is known, and this greatly simplifies the analysis. If the substrate is much thicker than the epitaxial layer, it is assumed to be unstrained ("thick substrate approximation"), with its normal Bragg angle θ_{BS}. It is most convenient to use symmetric rocking curves (i.e., 004 for zinc blende (001) semiconductors or 0002 for wurtzite (0001)) for the analysis.

The angular separation between the epitaxial layer and substrate peaks in the rocking curve (symmetric 00 *m* reflection) at an azimuth ψ is

$$\Delta\theta_{00m}(\psi) = \Delta\theta_{B00m} + \Delta\phi\cos(\psi - \psi_0), \tag{7.74}$$

where:

$\Delta\theta_{B00m}$ is the difference in 00 *m* Bragg angles between the epitaxial layer and the substrate, $\Delta\theta_{B00m} = \theta_{B00m,\text{epitaxial}} - \theta_{B00m,\text{substrate}}$

$\Delta\phi$ is the crystallographic tilt between the [001] axes of the epitaxial layer and the substrate

ψ_0 specifies the direction of this tilt. In order to find the strain in the epitaxial layer, it is necessary to measure rocking curves at two or more azimuths

If rocking curves are measured at opposing azimuths* $\psi = 0°$ and $\psi = 180°$, then

$$\Delta\theta_{B00m} = \frac{\Delta\theta(\psi = 0°) + \Delta\theta(\psi = 180°)}{2}. \tag{7.75}$$

The out-of-plane lattice constant is then determined from the Bragg angle for the epitaxial layer.

For a diamond or zinc blende heteroepitaxial layer, using the 00 *m* reflection,

$$c = \frac{m\lambda}{2\sin(\theta_{B00m,\text{substrate}} + \Delta\theta_{B00m})}. \tag{7.76}$$

The out-of-plane strain is

$$\varepsilon_\perp = \frac{c - a_0}{a_0}, \tag{7.77}$$

and assuming biaxial stress and tetragonal distortion, the in-plane strain is

$$\varepsilon_\parallel = -\frac{2C_{12}}{C_{11}}\varepsilon_\perp. \tag{7.78}$$

Similarly, for a wurtzite or hexagonal SiC crystal, using the 000 *m* reflection,

$$c = \frac{m\lambda}{2\sin(\theta_{B000m,\text{substrate}} + \Delta\theta_{B000m})}. \tag{7.79}$$

The out-of-plane strain is

$$\varepsilon_\perp = \frac{c - c_0}{c_0}, \tag{7.80}$$

and assuming biaxial stress, the in-plane strain is

$$\varepsilon_\parallel = -\frac{2C_{13}}{C_{33}}\varepsilon_\perp. \tag{7.81}$$

* The reference for the azimuth is arbitrary and can be set for convenience.

If it is necessary to determine the crystallographic tilt, then the rocking curve must be measured at least one more azimuth.

For the case of biaxial strain, the strain and stress may be determined readily from Raman spectroscopy measurements.[63] The shift of a given phonon mode is linearly related to the in-plane biaxial stress, which for a III-nitride semiconductor is given by

$$\Delta\Omega_\lambda = \frac{2C_{33}a_\lambda - 2b_\lambda C_{13}}{(C_{11}+C_{12})C_{33} - 2C_{13}^2}\sigma_{||}, \tag{7.82}$$

where the phonon deformation constants a_λ and b_λ have been tabulated by Davydov et al.[63] for several phonon modes in GaN.

7.8.2 X-Ray Diffraction Analysis of a Ternary Heteroepitaxial Layer

For a ternary alloy layer such as $In_xGa_{1-x}As$, or an alloy such as Ge_xSi_{1-x}, the independent determination of the relaxed lattice constant (and therefore the composition) and the state of strain requires measurements of two different *hkl* rocking curves. Sometimes, the analysis is simplified with the assumption that the heteroepitaxial layer has grown coherently on the substrate.[64–66] With this "pseudomorphic" assumption, the in-plane lattice constant is assumed to be equal to that of the substrate. Then a single rocking curve measurement, using a symmetric reflection, is used for the estimation of the composition and state of strain in a ternary layer. This simplified approach has been extended to quaternary semiconductors, for which a single x-ray rocking curve measurement is combined with a PL measurement to determine the bandgap (and therefore the composition and relaxed lattice constant) for the material. Such a simplified approach is suitable for a heteroepitaxial system such as AlGaAs/GaAs, for which the lattice mismatch strain is small over the entire range of composition. Usually, however, it is not possible to start with the pseudomorphic assumption.

Typically, for heteroepitaxy of a (001) zinc blende substrate, rocking curves are obtained for one symmetric reflection, such as the 004, and one asymmetric reflection, such as the 115 or 044. Then, with the assumption that the strained alloy is tetragonally distorted, the in-plane and out-of-plane lattice constants (*a* and *c*, respectively) may be determined. However, it is necessary to account for both the crystallographic tilting of the epitaxial layer with respect to the substrate and the additional tilting of the asymmetric planes due to the tetragonal distortion. The standard procedure for analysis of a zinc blende epitaxial layer using a symmetric 00m reflection and an asymmetric *hkl* reflection will be described below. However, the adaptation of this procedure to hexagonal epitaxial layers is a straightforward extension.

As in the case of the nonalloyed ("binary") semiconductor layer, the analysis starts with a symmetric reflection, which is measured at two opposing azimuths. If the 004 reflection is used, then the rocking curve peak separation at an azimuth ψ will be

$$\Delta\theta_{B004}(\psi) = \Delta\theta_{B004} + \Delta\phi\cos(\psi - \psi_0), \tag{7.83}$$

where:

- $\Delta\theta_{B004}$ is the difference in 004 Bragg angles between the epitaxial layer and the substrate
- $\Delta\phi$ is the crystallographic tilt between the epitaxial layer and the substrate
- ψ_0 specifies the direction of this tilt

The difference in Bragg angles for the 004 reflection is found by averaging the peak separation for two opposing azimuths,

$$\Delta\theta_{B004} = \frac{\Delta\theta(\psi = 0^\circ) + \Delta\theta(\psi = 180^\circ)}{2}. \tag{7.84}$$

The out-of-plane lattice constant can be determined as before, with the assumption that the substrate is unstrained ("thick substrate approximation"),

$$c = \frac{4\lambda}{2\sin(\theta_{B004,\text{substrate}} + \Delta\theta_{B004})}. \tag{7.85}$$

An additional complication arises if one attempts to use the above approach with an asymmetric reflection such as the 044. In such cases, biaxial strain in the heteroepitaxial layer gives rise to tetragonal distortion of the lattice, and so an additional tilt component $\Delta\phi_{\text{tet}}$ must be considered:

$$\Delta\theta_{B044}(\psi) = \Delta\theta_{B044} + \Delta\phi\cos(\psi - \psi_0) + \Delta\phi_{\text{tet}}(044), \tag{7.86}$$

where $\Delta\phi_{\text{tet}}(044)$ is given by

$$\Delta\phi_{\text{tet}}(044) = \cos^{-1}\left(\frac{4}{c\sqrt{\left(\frac{4}{a}\right)^2 + \left(\frac{4}{c}\right)^2}}\right) - \cos^{-1}\left(\frac{4}{\sqrt{4^2 + 4^2}}\right). \tag{7.87}$$

More generally, in the case of an *hkl* reflection from a substrate with (001) orientation, the angle between the (*hkl*) planes and the surface is given by

$$\phi(hkl) = \cos^{-1}\left(\frac{l}{\sqrt{h^2 + k^2 + l^2}}\right). \tag{7.88}$$

The change in ϕ due to tetragonal distortion is then

$$\Delta\phi_{\text{tet}}(hkl) = \cos^{-1}\left(\frac{l}{c\sqrt{\left(\frac{h}{a}\right)^2 + \left(\frac{k}{a}\right)^2 + \left(\frac{l}{c}\right)^2}}\right) - \cos^{-1}\left(\frac{l}{\sqrt{h^2 + k^2 + l^2}}\right). \tag{7.89}$$

Like before, the measurement of the asymmetric rocking curves at opposing azimuths, *for the same set of planes*, allows elimination of the tilt component $\Delta\phi_{\text{tet}}(hkl)$.[67–69] However, the disadvantage of that approach is that it requires measuring the rocking curve for one azimuth using $\theta_B - \phi$ incidence, as shown in Figure 7.29. This leads to a rocking curve peak that is broadened and weakened in intensity. Specifically, the intensity ratio for the two angles of incidence can be estimated as[70]

$$\frac{I(\theta_B + \phi)}{I(\theta_B - \phi)} = \frac{\sin^2(\theta_B + \phi)}{\sin^2(\theta_B - \phi)}, \tag{7.90}$$

where $I(\theta_B + \phi)$ and $I(\theta_B - \phi)$ are the intensities for $\theta_B + \phi$ and $\theta_B - \phi$ incidence, respectively. For example, in the case of the 044 reflection from GaAs(001) with Cu kα radiation, the

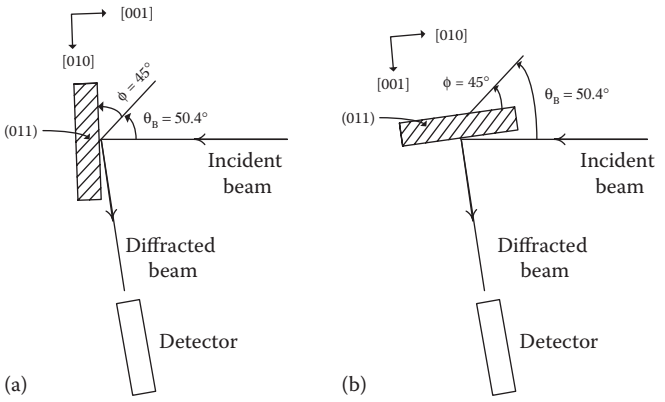


FIGURE 7.29
Asymmetric 044 reflections from a (001) zinc blende crystal, at opposing azimuths using the same set of diffracting planes. (a) $\psi = 0°$ with $\theta_B + \phi$ incidence, and (b) $\psi = 180°$ with $\theta_B - \phi$ incidence. (Adapted from X. G. Zhang et al., *J. Vac. Sci. Technol. B*, 18, 1375 [2000]. With permission. Copyright 2000, American Vacuum Society.)

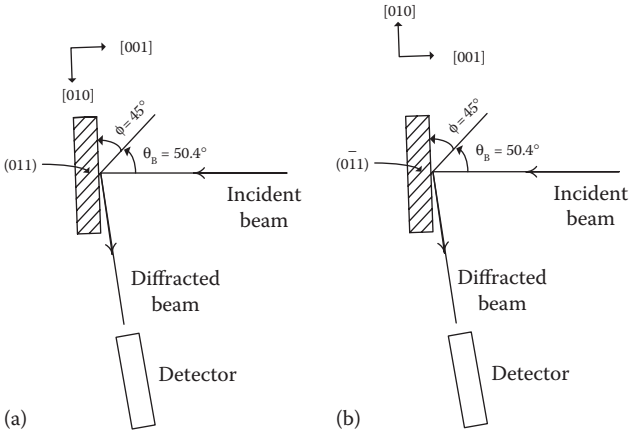


FIGURE 7.30
Asymmetric 044 reflections from a (001) zinc blende crystal, at opposing azimuths using two different sets of diffracting planes. (a) $\psi = 0°$ and (b) $\psi = 180°$. Both rocking curves are measured at $\theta_B + \phi$ incidence. (Adapted from X. G. Zhang et al., *J. Vac. Sci. Technol. B*, 18, 1375 [2000]. With permission. Copyright 2000, American Vacuum Society.)

intensity ratio is 112. This means that the reflected intensity will be insufficient for the purpose of an accurate measurement in that case. Instead, it is necessary to measure the rocking curves at both azimuths with the $\theta_B + \phi$ incidence, as shown in Figure 7.30. Then the two rocking curves are measured from two different sets of {011} planes. The tilt $\Delta\phi_{tet}$ does not cancel out when we average the two peak separations, but must be extracted in the analysis.

Using an asymmetric *hkl* reflection, the peak separation between the epitaxial layer and the substrate is averaged for the two opposing azimuths, with both rocking curves measured at $\theta_B + \phi$ incidence. For example, if ψ_0 corresponds to the azimuth of the first measurement, then the average peak separation is

$$\Delta\theta_{\mathrm{AVE},hkl} = \frac{\Delta\theta(\psi_0) + \Delta\theta(\psi_0 + 180°)}{2}. \tag{7.91}$$

For example, using the 044 reflection, the two measurements are taken at $\psi = 45°$ and $\psi = 225°$. It should be noted that $\Delta\theta_{\mathrm{AVE},hkl}$ found in this way is not the same as the Bragg angle difference, but includes a tilt component due to the tetragonal distortion of the crystal, $\Delta\phi_{\mathrm{tet}}$. The spacing of the *hkl* planes can be found from

$$d_{hkl} = \frac{\lambda}{2\sin\left(\theta_{Bhkl,\mathrm{substrate}} + \Delta\theta_{\mathrm{AVE},hkl} - \Delta\phi_{\mathrm{tet}}\right)}. \tag{7.92}$$

The in-plane lattice constant in the strained heteroepitaxial layer is found from

$$a = \left(\frac{h^2 + k^2}{l^2/c^2 - l/d_{hkl}^2}\right)^{-1/2}. \tag{7.93}$$

The tilting of the *hkl* planes due to the tetragonal distortion can be calculated as

$$\Delta\phi_{\mathrm{tet}}(hkl) = \cos^{-1}\left(\frac{l/c}{\sqrt{(h/a)^2 + (k/a)^2 + (l/c)^2}}\right) - \cos^{-1}\left(\frac{l}{\sqrt{h^2 + k^2 + l^2}}\right). \tag{7.94}$$

Equations 7.92 through 7.94 must be solved iteratively, starting with an assumed value of $\Delta\phi_{\mathrm{tet}}$. Typically, the values of d_{hkl}, a, and $\Delta\phi_{\mathrm{tet}}$ converge after six or fewer iterations.

Once the in-plane and out-of-plane lattice constants have been determined, the relaxed lattice constant may be calculated from

$$a_0 = \frac{c + a\left(\frac{2\nu}{1-\nu}\right)}{1 + \left(\frac{2\nu}{1-\nu}\right)}, \tag{7.95}$$

where ν is the Poisson ratio. The relaxed lattice constant is used to determine the composition, either by using the known lattice constant versus composition characteristic or by using a linear interpolation with the lattice constants of the binaries (Vegard's law). If the Poisson ratio varies strongly with composition, iteration can be used to find a consistent solution for a_0 and ν. Finally, the in-plane and out-of-plane strains in the heteroepitaxial layer are found using

$$\varepsilon_{\parallel} = \frac{a - a_0}{a_0} \tag{7.96}$$

and

$$\varepsilon_{\perp} = \frac{c - a_0}{a_0}. \tag{7.97}$$

7.8.3 X-Ray Analysis of a Quaternary Heteroepitaxial Layer

In the case of a quaternary alloy such as $Al_xIn_yGa_{1-x-y}P$, the composition is not uniquely determined once the relaxed lattice constant is known. For this reason, the analysis of the composition and strain cannot be done using x-ray rocking curves alone. If, however, the composition (and therefore the relaxed lattice constant) is determined by some other technique, such as auger electron spectroscopy (AES) or secondary ion mass spectroscopy (SIMS), the in-plane and out-of-plane strains may be determined by following the procedure above. Another approach involves the estimation of the bandgap from PL measurements. The composition is not uniquely determined by the bandgap. Nonetheless, the knowledge of the bandgap (from PL) and the relaxed lattice constant (from XRD) together allows the determination of the composition. The in-plane and out-of-plane strains are then determined as before from the XRD values of a, c, and a_0. An important source of error in such a procedure is the strain-induced shift in the PL emission peak. Therefore, the XRD strain results must be used to estimate this shift if reliable results are to be obtained.

7.8.4 Analysis of Composition and Strain Using X-Ray Reciprocal Space Maps

X-ray RSMs contain a great wealth of information and may be used to determine strains, compositions, crystallographic tilts, and defect densities in semiconductor structures. The application to uniform or linearly graded structures is fairly straightforward, as will be illustrated by the examples given below. On the other hand, the characterization of multilayer device structures is fairly complex, and full utilization of the potential offered by mapping may only be achieved by the full development of dynamical RSM simulations.

7.8.4.1 Application of RSM to a Uniform Buffer Layer

Fewster and Andrew[41] measured RSMs for a 6 μm thick $In_{0.05}Ga_{0.95}As$ buffer layer grown on GaAs(001) by MBE. This sample also had an $In_{0.10}Ga_{0.90}As$/GaAs superlattice grown on top, as will be discussed below. The RSMs were measured in the vicinity of the 004 and 115 reciprocal lattice points, in each case for two azimuths $\pi/2$ apart, and the resulting maps are shown in Figure 7.31. Here, the intensity is indicated by the density of the dark pixels, so that intensity contours have not been shown. Relative displacements $\Delta q_\perp$ and $\Delta q_\parallel$ have been given; these are expressed as $\Delta q_{[004]}$ and $\Delta q_{[110]}$ for the 004 reciprocal lattice point, and $\Delta q_{[115]}$ and $\Delta q_{[552]}$ for the 115 reciprocal lattice point, respectively. Rotation of the azimuth by $\pi/2$ renders the $\Delta q_\parallel$ to $\Delta q_{[1\bar{1}0]}$ and $\Delta q_{[5\bar{5}2]}$ for the 004 and 115 reflections, respectively. From the 004 maps, the displacement in $\Delta q_\perp$ allows determination of the out-of-plane lattice constant for the $In_{0.5}Ga_{0.95}As$ layer, whereas a tilting of the $In_{0.05}Ga_{0.95}As$ layer shows up as a displacement in $\Delta q_\parallel$. Therefore, from the two 004 RSMs measured at orthogonal azimuths, it is possible to independently determine the components of tilt about the [110] and $[1\bar{1}0]$ axes, by translating back to the angular scale. The 115 RSMs are influenced by the in-plane lattice constant for the $In_{0.5}Ga_{0.95}As$ buffer, as well as its tilt, but these effects can be separated to enable the determination of all three lattice constants for the strained buffer layer with orthorhombic symmetry. From an analysis of the RSMs, it was determined that the lattice relaxation in the $In_{0.5}Ga_{0.95}As$ buffer was 108%; in other words, the in-plane lattice parameter shrank below its cubic value, perhaps as a consequence of thermal strain or the influence of tilts on the relaxation process. The zeroth-order superlattice peak and its satellites are "swamped out" by the intensity of the buffer, but lie along the surface normal direction, indicating that the superlattice is coherently grown on the relaxed buffer layer.

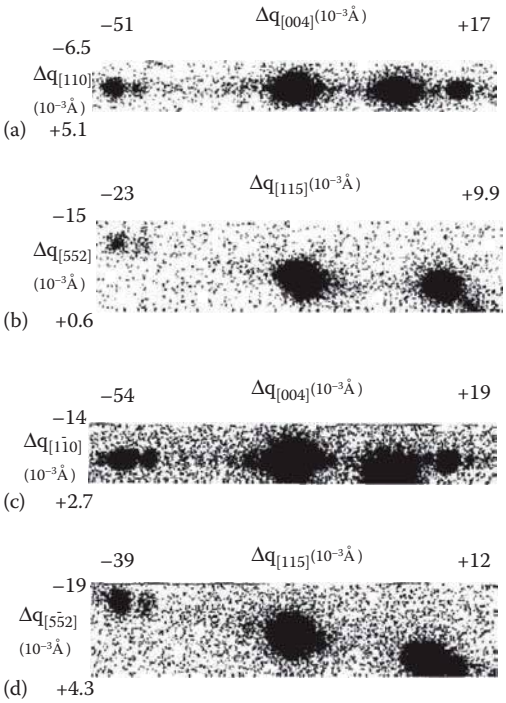


FIGURE 7.31
RSMs for a 6 μm thick $In_{0.05}Ga_{0.95}As$ buffer on a GaAs(001) substrate. Maps (a) and (c) were measured in the vicinity of the 004 reciprocal lattice point at two azimuths separated by π/2. Maps (b) and (d) were measured in the vicinity of the 115 reciprocal lattice point at two azimuths separated by π/2. (Reprinted from P. F. Fewster and N. L. Andrew, *J. Appl. Phys.*, 74, 3121 [1993]. With permission. Copyright 1993, American Institute of Physics.)

Koppensteiner et al.[71] measured RSMs for an MBE-grown single-step 20 nm $Si_{0.6}Ge_{0.4}$ buffer on a Si(001) substrate, with a short-period Si_6Ge_4 superlattice (145 periods) on top. The RSMs were measured in the vicinity of the 004 and 224 reciprocal space points and are shown in Figure 7.32. Here, the maps are shown using isointensity contours that correspond to 0.5, 1, 2, 10, 20, and 2000 counts per second in the 224 map, and 1, 2, 10, 20, 100, and 2000 counts per second in the 004 map. The asymmetry seen in the 004 substrate reciprocal space point is an artifact caused by the finite resolution of the analyzer, which in this case was a slit instead of an analyzer crystal. These maps show that the buffer was partially relaxed, and that the short-period superlattice was grown coherently on the buffer.

Using a symmetric reflection, such as the 004 reflection from a (001)-oriented sample, the out-of-plane lattice constant of a uniform epitaxial layer may be determined by[72]

$$\left(\frac{c-a_s}{a_s}\right) = -\frac{\Delta q_\perp}{h_{s,\perp}+\Delta q_\perp}, \tag{7.98}$$

where:

$h_{s,\perp}$ is the out-of-plane component of the substrate reciprocal lattice vector
a_s is the substrate lattice constant
$\Delta q_\perp$ is the RSM displacements from the substrate reciprocal lattice point

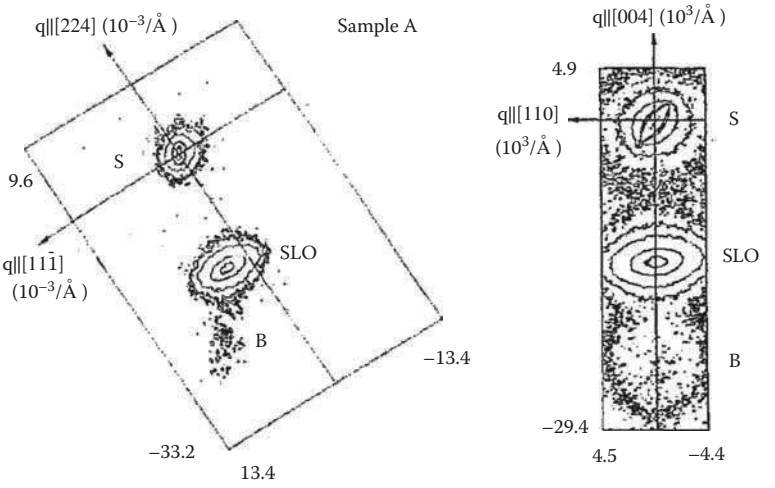


FIGURE 7.32
224 and 004 RSMs for an MBE-grown 20 nm $Si_{0.6}Ge_{0.4}$ buffer on a Si(001) substrate, with a short-period Si_6Ge_4 superlattice (145 periods) on top. The features marked S, SLO, and B correspond to the substrate, superlattice zeroth-order peak, and buffer layer, respectively. The isointensity contours correspond to 0.5, 1, 2, 10, 20, and 2000 counts per second in the 224 map, and 1, 2, 10, 20, 100, and 2000 counts per second in the 004 map. (Reprinted from E. Koppensteiner et al., *Appl. Phys. Lett.*, 62, 1783 [1993]. With permission. Copyright 1993, American Institute of Physics.)

It should be noted here that $\Delta q_\perp$ must be corrected to remove any crystallographic tilting of the epitaxial layer with respect to the substrate.

Using an asymmetric reflection, such as the 224, 026, 135, or 115, it is possible to determine both the in-plane and out-of-plane lattice constants a and c of a uniform layer from the RSM as follows:[72]

$$\left(\frac{a-a_s}{a_s}\right) = -\frac{\Delta q_{\|}}{h_{s,\|} + \Delta q_{\|}} \tag{7.99}$$

and

$$\left(\frac{c-a_s}{a_s}\right) = -\frac{\Delta q_{\perp}}{h_{s,\perp} + \Delta q_{\perp}}, \tag{7.100}$$

where:

$h_{s\prime\|}$ and $h_{s\prime\perp}$	are the in-plane and out-of-plane components of the substrate reciprocal lattice vector
a_s	is the substrate lattice constant
$\Delta q_{\|}$ and $\Delta q_{\perp}$	are the RSM displacements from the substrate reciprocal lattice point, corrected for any crystallographic tilting of the *hkl* planes arising from crystallographic tilting or tetragonal distortion

The tilt of a uniform layer relative to the (001) substrate may be determined from the 004 RSM by[73]

$$\phi = \tan^{-1}\left[\frac{\left|\Delta q_{\parallel,0} - \Delta q_{\parallel,180}\right|/2}{4/a_s - \left|\Delta q_{\perp}\right|}\right], \tag{7.101}$$

where:

$\Delta q_{\parallel,0}$ is the reciprocal space displacements in the in-plane direction measured at 0°

$\Delta q_{\parallel,180}$ is the reciprocal space displacements in the in-plane direction measured at 180° azimuth

$\Delta q_{\perp}$ is the reciprocal space displacement in the out-of-plane direction, and where the tilt is about the line in the surface corresponding to the 90° azimuth

(Usually, the axis of the crystallographic tilt is not known *a priori*, so it may be necessary to measure RSMs at four azimuths, 0°, 90°, 180°, and 270°, in order to first establish the axis of the tilt.)

7.8.4.2 Application of RSM to Linearly Graded and Step-Graded Buffers

RSMs may also be used to assess the state of strain and crystallographic tilting in linearly graded and step-graded buffers. As with the analysis of HRXRD rocking curves, the tilting is easily determined by using a symmetric reflection measured at two or more azimuths. Once the tilt is accounted for, the RSM shows the variations of the in-plane and out-of-plane lattice constants in the graded structure. For example, Figure 7.33 shows RSMs for LED structures on uniform 1 μm thick $InAs_ySb_{1-y}$ grown on GaSb using an intermediate linearly graded buffer of $InAs_xSb_{1-x}$ and measured around the 335 reflection of the GaSb(001) substrate.[74] The mole fractions of Sb in the top layer were 0.2 (a and b) and 0.44 (c and d). RSMs were measured at azimuths of 0° and 180° with respect to the [110] direction to allow cancellation of the crystallographic tilt. The abscissa and ordinate are $3/2a_{\parallel}$ and $5/a_{\perp}$ respectively, where $a_{\parallel}$ and $a_{\perp}$ are the in-plane and out-of-plane lattice constants, respectively. The red line shows the loci of 335 reciprocal lattice points from fully relaxed cubic material with varying lattice constants. The perpendicular blue lines represent a constant in-plane lattice constant, coherently grown with the top of the linearly graded buffer. In each of the two structures, the linearly graded buffer gives rise to intensity along the diagonal line, indicating nearly complete relaxation. For the structure with 20% antimony, the uniform layer reciprocal lattice point, indicated by L, is off the diagonal line. From its position, the in-plane and out-of-plane lattice constants may be determined, and if the elastic constants are estimated, the relaxed lattice constant and state of strain may be determined.

Jiang et al. measured RSMs around the 004 and 224 reciprocal lattice points for a 900 nm overshoot step-graded $In_xAl_{1-x}As$ buffer with nine steps from x = 0.05 to x = 0.85, and a 1 μm uniform $In_{0.75}Al_{0.25}As$ top layer grown on a GaAs(001) substrate, and these are shown in Figure 7.34. The substrate peak is centered at the origin, while the uniform $In_{0.75}Al_{0.25}As$ peak is located at the bottom (004) or bottom left (224) of the map. The intensity from the step-graded buffer is spread between these two, and even extends beyond the peak of the uniform layer, because of the compositional overshoot built into the buffer layer. The buffer layer intensity lies on a line corresponding to completely relaxed material, which is vertical for the 004 map and inclined for the 224 map. In this sample, the intensity from the step-graded buffer seems to span a continuum because of the large number of steps. Discrete peaks are often seen in the RSM for a step-graded buffer, as shown in Figure 7.35a (top two maps) for an $In_xAl_{1-x}As$/GaAs(001) step-graded buffer grown by MBE.[72]

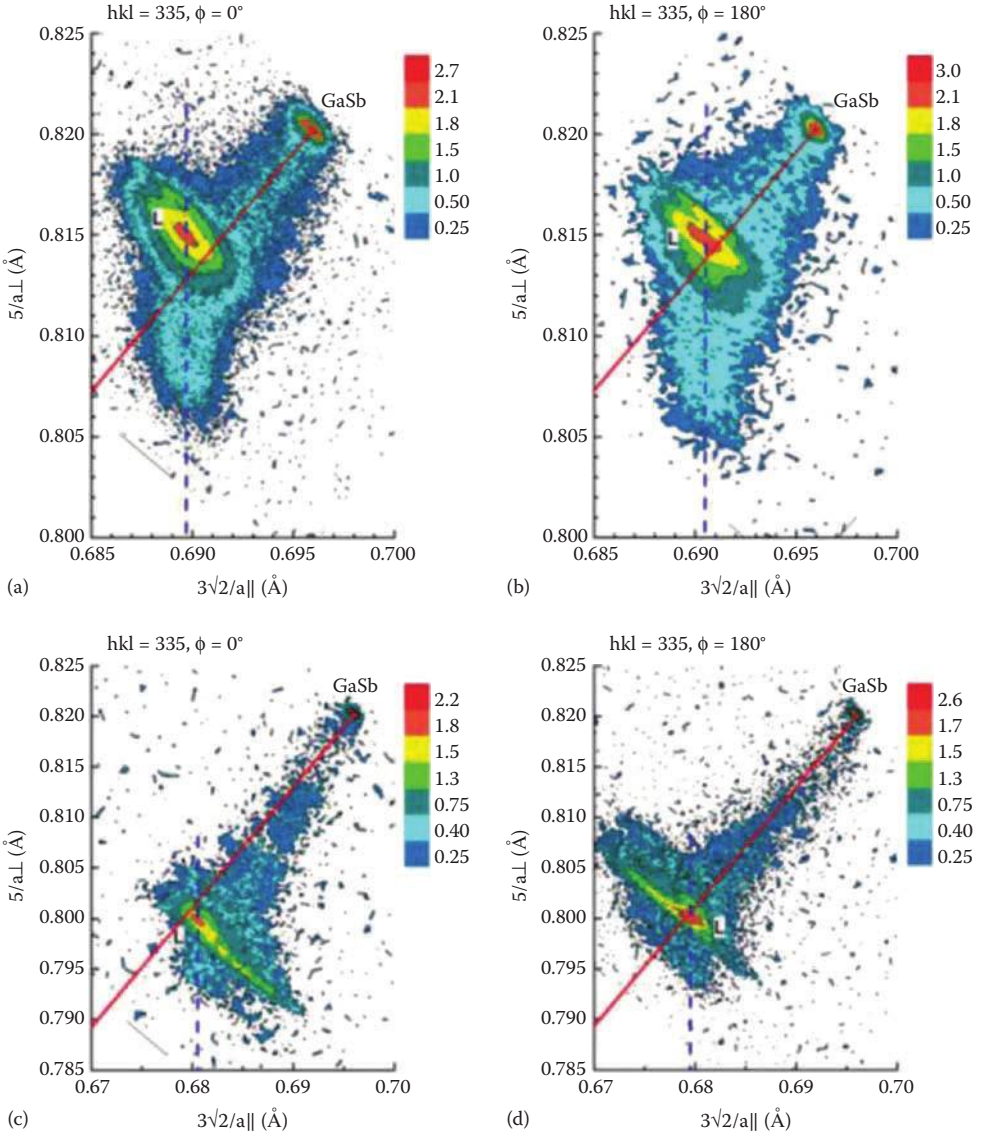


FIGURE 7.33
RSMs for LED structures on uniform 1 μm thick $InAs_ySb_{1-y}$ grown on GaSb using an intermediate linearly graded buffer of $InAs_xSb_{1-x}$ and measured around the 335 reflection of the GaSb(001) substrate. The mole fractions of Sb in the top layer were 0.2 (a, b) and 0.44 (c, d). (Reprinted from G. Belenky et al., *Appl. Phys. Lett.*, 102, 111108 [2013]. With permission. Copyright 2013, American Institute of Physics.)

7.8.4.3 Application of Reciprocal Space Maps to III-Nitride Materials

Pereira et al.[75,76] have discussed the application of RSMs to III-nitride epitaxial layers. Figure 7.36 shows the loci of reciprocal lattice points for InGaN layers on GaN. In the case of a symmetric reflection such as the 0002 (also referred to as the 002), the reciprocal lattice point of the epitaxial layer will be located on a vertical line passing through the GaN reciprocal lattice point, regardless of the state of strain in the layer. In this case, the vertical displacement

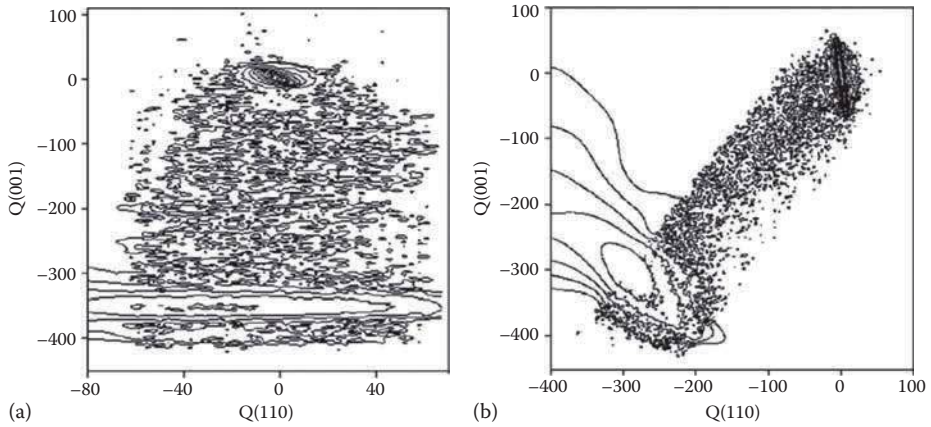


FIGURE 7.34
RSMs measured around the (a) 004 and (b) 224 reciprocal lattice points for a 900 nm overshoot step-graded $In_xAl_{1-x}As$ buffer with nine steps from x = 0.05 to x = 0.85, and a 1 μm uniform $In_{0.75}Al_{0.25}As$ top layer grown on a GaAs(001) substrate. The substrate peak is centered at the origin, while the uniform $In_{0.75}Al_{0.25}As$ peak is located at the bottom (004) or bottom left (224) of the map. (Reprinted from Z. Jiang et al., *Appl. Surf. Sci.*, 254, 5241 [2008]. With permission. Copyright 2008, Elsevier.)

is determined by both the strain and composition. For an asymmetric reflection such as the 10–15 (also referred to as the 105), fully strained (coherent) epitaxial layers will exhibit reciprocal lattice points on a vertical line passing through the GaN reciprocal lattice point, but fully relaxed layers will have reciprocal lattice points on a diagonal line, as shown in Figure 7.36. Therefore, the measurement of one symmetric reflection and one asymmetric reflection allows the independent determination of the strain and composition in the absence of crystallographic tilt. If tilt is present, it can be determined by measuring symmetric reflections at two or more azimuths, and then the RSMs may be corrected for the tilt. Figure 7.37 shows 105 RSMs for InGaN/GaN(0001) layers of three different thicknesses and illustrates the application of these concepts. The 60 nm sample (sample A) exhibits a reciprocal lattice point that is close to the vertical line, showing that this sample is nearly coherently strained. With increasing thickness in sample B (120 nm) and sample C (240 nm), the peak intensity of the epitaxial layer moves toward the fully relaxed line. However, a splitting of the intensity in two peaks shows that some of the material remains coherent, while some of the material relaxes. This is also shown in the 105 RSM of Figure 7.38 for a 120 nm layer of $In_{0.19}Ga_{0.81}N$/GaN that exhibits both coherently strained (S) and relaxed (R) reciprocal lattice points. The streaking of the GaN reciprocal space point in this map is an artifact caused by the instrumental effects (wavelength spread and finite acceptance angle of the analyzer).

Although most RSM studies of III-nitrides have been conducted with *c*-plane material, Moram et al.[77] have discussed the application of reciprocal space mapping to nonpolar *a*-plane GaN films. Although the strain analysis is similar to the *c*-plane case, Moram et al. show that the broadening of the reciprocal space points is related to a number of factors, including wafer curvature, surface roughness, and stacking faults, as well as threading dislocations, making quantitative analysis difficult.

7.8.5 *In Situ* Stress–Strain Measurements Using Multibeam Optical Stress Sensor

It is possible to study the evolution of the stress in a heteroepitaxial film in the growth environment using a multibeam optical stress sensor (MOSS). The deposition of a biaxially

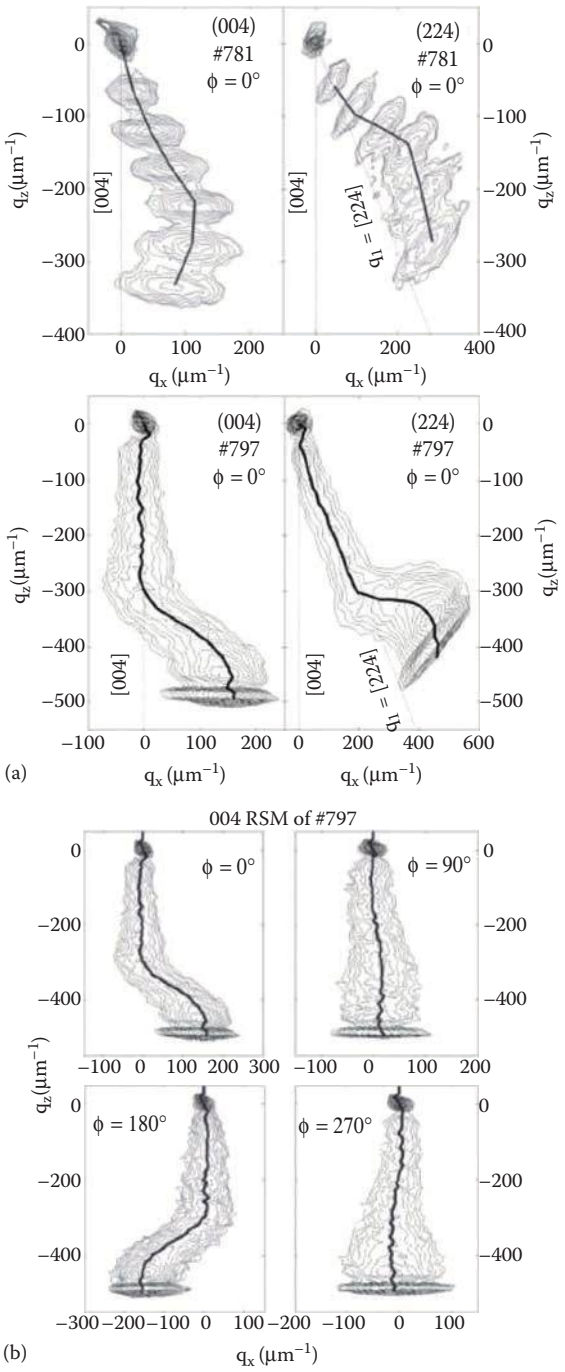


FIGURE 7.35
(a) RSMs for a step-graded $In_xAl_{1-x}As$/GaAs(001) buffer and a linearly graded $In_xAl_{1-x}As$/GaAs (001) buffer layer measured in the vicinity of the 004 and 024 reciprocal lattice points with zero azimuthal angle. (b) RSMs for a linearly graded $In_xAl_{1-x}As$/GaAs(001) buffer measured in the vicinity of the 004 reciprocal lattice point for four different azimuthal angles. (Reprinted from D. Lee et al., *J. Appl. Phys.*, 101, 063523 [2007]. With permission. Copyright 2007, American Institute of Physics.)

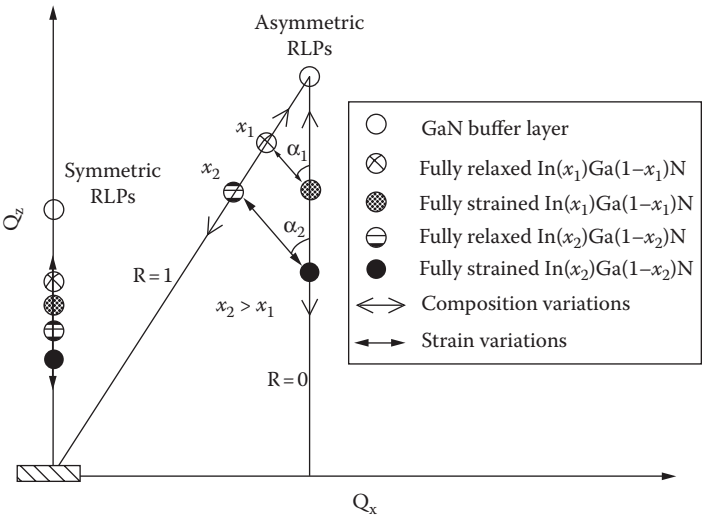


FIGURE 7.36
Loci of reciprocal lattice points for InGaN layers that are fully strained or fully relaxed, for symmetric and asymmetric x-ray reflections. (Reprinted from S. Pereira et al., *Appl. Phys. Lett.*, 80, 3913 [2002]. With permission. Copyright 2002, American Institute of Physics.)

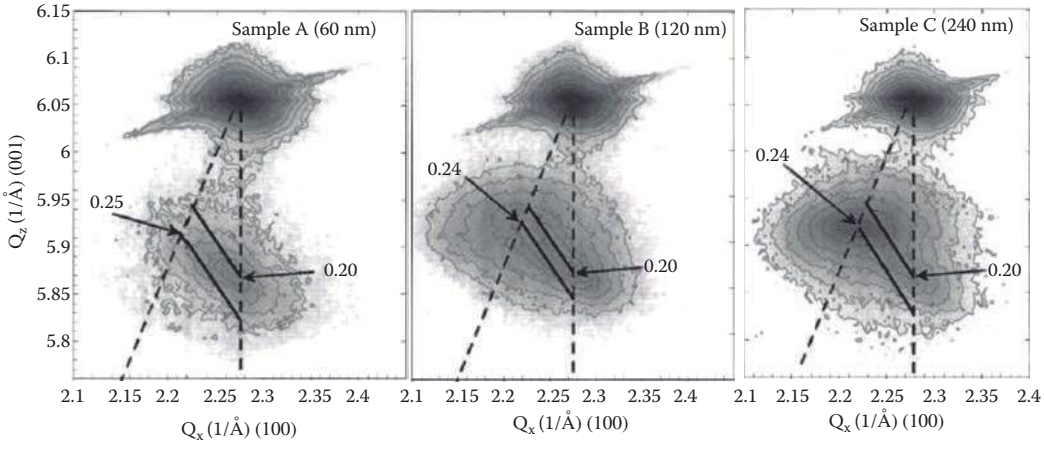


FIGURE 7.37
RSMs measured in the vicinity of the 105 reflection for InGaN/GaN layers with three different thicknesses of 60 nm (sample A), 120 nm (sample B), and 240 nm (sample C). (Reprinted from S. Pereira et al., *Appl. Phys. Lett.*, 80, 3913 [2002]. With permission. Copyright 2002, American Institute of Physics.)

strained film exerts forces on the substrate that are translated into an in-plane stress, as well as a bending of the substrate. For thin films that are much thinner than the substrate, Stoney[78] derived the expression for the substrate curvature κ in relation to the mean biaxial film stress σ_f given by

$$\kappa = \frac{1}{R} = \frac{6\sigma_f h_f}{Y_s h_s^2}, \tag{7.102}$$

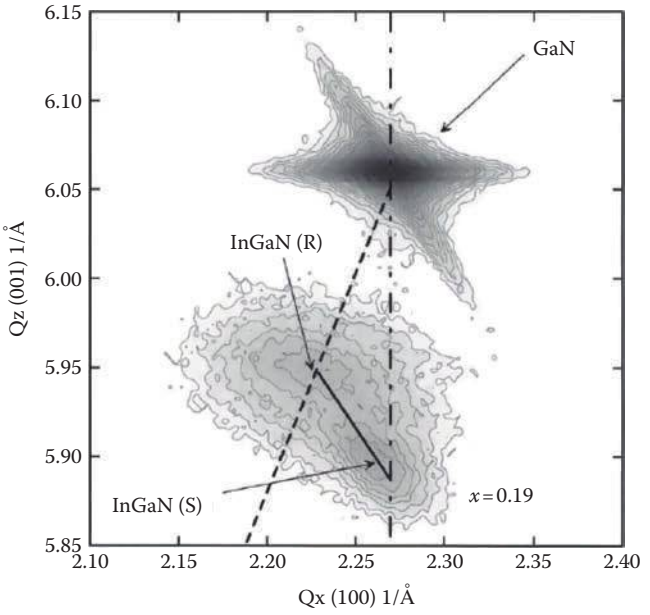


FIGURE 7.38
105 RSM for 120 nm $In_{0.19}Ga_{0.81}N$/GaN(0001). The streaks seen in the GaN reciprocal space point are artifacts associated with the instrumental effects. (Reprinted from S. Pereira et al., *Appl. Phys. Lett.*, 81, 1207 [2002]. With permission. Copyright 2002, American Institute of Physics.)

where:

R	is the radius of curvature
h_f and h_s	are the epilayer and substrate thickness
Y_s	is the biaxial modulus of the substrate

Equation 7.102 assumes the epilayer has a flat surface and both the epilayer and substrate are completely isotropic and free of defects. A large body of work has been published on the application of Stoney's original form, and the new extensions deal with compositionally graded layers,[79] anisotropic epilayer–substrate systems and nonlinear deformation regimes,[80] and systems with large deformations.[81] This work has enabled the demonstration of real-time measurements of the stress and relaxation behavior with respect to two-dimensional growth[82] and islanding,[83,84] in the SiGe/Si[82–84] and InGaAs[85–89] material systems. Figure 7.39 shows a simple experimental setup adapted from the works of Floro et al.[82] and Beresford et al.,[89] which discuss the use of the MOSS. Based on Stoney's equation and the geometrical parametrization of the optical beam path shown in Figure 7.39, the substrate curvature is given by

$$\kappa = \frac{1}{R} = \frac{6\sigma_f h_f}{Y_s h_s^2} = \left(\frac{\delta d}{d}\right)\frac{\cos\alpha}{2L}, \tag{7.103}$$

where:

α	is the angle of incidence
$\delta d/d$	is the mean differential spacing (MDS)
L	is the optical path
d	is the laser beam spacing

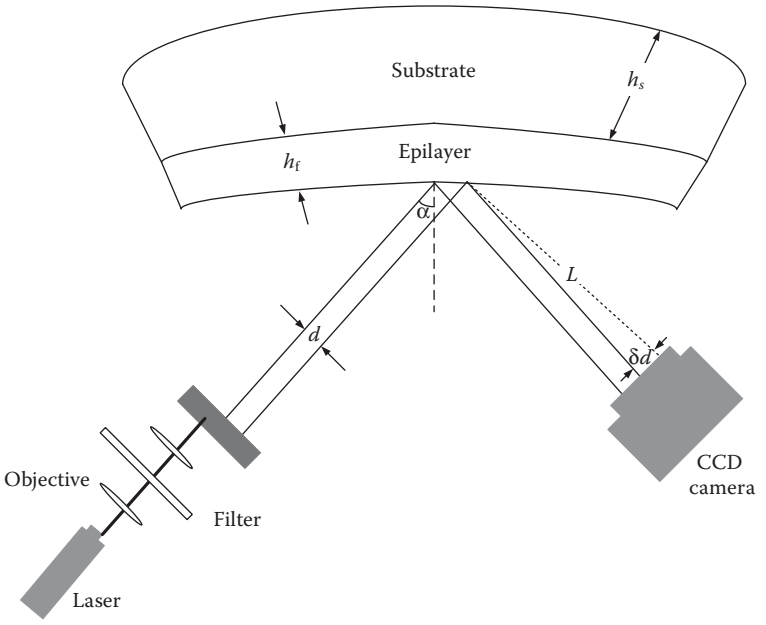


FIGURE 7.39
Schematic representation of the optical stress sensor measurement setup.

The substrate curvature is determined from measurements of the MDS, $\delta d/d$, of the reflected spots. The growth of the InGaAs is monitored by the evolution of the MDS; as the growth proceeds, the MDS experiences a linear increase that is associated with the region of pseudomorphic growth, and in this region of constant stress, the MDS is proportional to epilayer thickness (h_f) and thus $\delta d/d \propto h_f$.

At the onset of relaxation, the MDS deviates from the linear increase and the stress in the film may be determined from the observed slope of the constant stress region (in other words, the pseudomorphic regime). The stress associated with a mismatch strain ε is given by

$$\sigma = Y_f \varepsilon = \left(\frac{\delta d/d}{t} \right) \frac{Y_s h_s^2 \cos\alpha}{12Lr}, \tag{7.104}$$

where:

t	is the growth time
$r = h_{f/t}$	is the growth rate
d	is the optical beam spacing
α	is the angle of incidence
L	is the optical path length

In the pseudomorphic region ($\varepsilon = f_0$), Equation 7.104 reduces to

$$\sigma_0 = Y_f f_0 = (s_i) \frac{Y_s h_s^2 \cos\alpha}{12Lr}, \tag{7.105}$$

where $s_i = (\delta d/d)/t$ is the initial slope in the constant stress region. The lattice relaxation ε at a particular point in time can be determined from

$$\gamma = \frac{f_0 - \varepsilon}{f_0} = 1 - \left(\frac{\delta d/d}{s_i t} \right). \tag{7.106}$$

Equation 7.106 shows that the measured *in situ* relaxation depends only on growth time and the final value of the MDS. It should be mentioned that the determined relaxation above from curvature measurements reflects the average biaxial stress state of the epilayer where

$$\langle \sigma \rangle = \frac{1}{h_f} \int_0^{h_f} \sigma(z) dz. \tag{7.107}$$

To obtain the biaxial stress at any point away from the substrate interface, we can differentiate Equation 7.107 with respect to time and implying that $h_f = rt$; then,

$$\sigma(z) = \langle \sigma \rangle + t \frac{d\langle \sigma \rangle}{dt}. \tag{7.108}$$

7.9 Determination of the Critical Layer Thickness

Experimental methods for the determination of the CLT include TEM,[90,91] scanning tunneling microscopy (STM),[92] PL,[90,93,94] PLM,[61,95] electrical measurements on modulation-doped structures,[96,97] RHEED,[93,98] XRD,[90,91,98,99] x-ray topography (XRT),[100] electron beam–induced current (EBIC),[101] and ion channeling.[102] The traditional strain-based methods for measuring h_c involve the determination of in-plane strain by PL or XRD. On the other hand, with imaging techniques such as TEM, PL microscopy, or XRT, individual dislocations (or the material surrounding them) can be imaged after the onset of lattice relaxation. There are also several methods for the indirect observation of lattice relaxation. For example, lattice relaxation introduces surface steps that broaden specular spots measured by RHEED. Also, the introduction of threading dislocations during lattice relaxation causes several observable changes in the XRD profiles. These include the broadening of the main diffraction peak, the broadening and extinguishing of the Pendellosung fringes,[91,100] and the reduction in the ratio of the epitaxial layer peak intensity to the substrate peak intensity. All of these phenomena have been used for the experimental determination of the critical thickness in mismatched heteroepitaxial layers.

In nearly all of the methods above, samples of various thicknesses are examined for evidence of strain relaxation. Therefore, thickness resolution (stemming from the use of a finite number of samples) is an important source of error unless many samples are grown and characterized.

Strain resolution is another source of error, which could be important if coupled with sluggish strain relaxation. Strain-based methods such as PL and XRD typically have strain resolutions of 10^{-5} to 10^{-4}. The resolution of an imaging technique such as TEM or PLM, though harder to quantify, can be better than 10^{-5} if a sufficiently large area is examined. All of these techniques appear to have sufficient resolution for the determination of the CLT. Despite this, anomalously large CLTs have sometimes been reported in studies based on XRD. New models for the CLT have been proposed to explain these results, for

example, by People and Bean[103] and Fischer et al.[104] It has also been shown by Fritz[105] that the anomalously large CLTs could be explained by initially sluggish lattice relaxation combined with finite experimental resolution. On the other hand, the anomalously large CLTs reported in XRD studies could be due to errors introduced by the crystallographic tilting of the epitaxial layers.[106] Fortunately, this cause for error can be readily eliminated.[107,108]

The choice of experimental technique will therefore be dictated in large part by the application. In the case of strained quantum wells for optoelectronic devices, PLM may be the most appropriate technique. If, however, strained layers are to be used in high-electron-mobility transistors (HEMTs), then measurements of the carrier mobility in modulation-doped structures are indicated. The study of strain relaxation kinetics can be best done using XRD. Finally, the types of dislocations and the mechanisms for their introduction are best studied using TEM.

7.9.1 Effect of Finite Resolution on CLT Determination

Fritz[105] has shown that limited strain resolution, coupled with initially sluggish lattice relaxation, can lead to the experimental determination of apparent CLTs that are much greater than the actual values. Essentially, any experimental method for CLT determination allows the measurement of the in-plane strain (directly or indirectly) for samples of various thickness. According to the Matthews and Blakeslee[109] model, the equilibrium strain in the layer is given by

$$\varepsilon_{\parallel}(\text{eq}) = \begin{cases} f; & h \le h_c \\ \dfrac{f}{|f|}\dfrac{b\left(1-\nu\cos^2\alpha\right)\left[\ln\left(h/b\right)+1\right]}{8\pi h\left(1+\nu\right)\cos\lambda}; & h > h_c \end{cases} . \tag{7.109}$$

The CLT is considered to correspond to the smallest thickness in which the measured strain departs measurably from the mismatch strain f.

Now, suppose the experimental method has finite resolution and can detect a change in strain no smaller than R. The CLT h_c^* determined with this finite experimental resolution will satisfy

$$|f| - R = \frac{b\left(1-\nu\cos^2\alpha\right)\left[\ln\left(h_c^*/b\right)+1\right]}{8\pi h_c^*\left(1+\nu\right)\cos\lambda} . \tag{7.110}$$

By way of example, the resolution of XRD techniques falls typically in the range 10^{-5} to 10^{-4}. Figure 7.40 shows the Matthews and Blakeslee[109] CLT h_c versus the lattice mismatch strain f, and also the experimental CLT h_c^*, assuming a worst case of $R = 10^{-4}$. So, although experimentally determined CLTs sometimes exceed the predictions of the Matthews and Blakeslee model significantly, this cannot be explained as a consequence of finite experimental resolution acting alone.

Partially relaxed layers with $h > h_c$ are usually found to contain fewer misfit dislocations than expected according to the equilibrium theory. If the ratio of the actual misfit dislocation density to the expected density for a layer in equilibrium is Q, then the apparent CLT h_c^* will satisfy

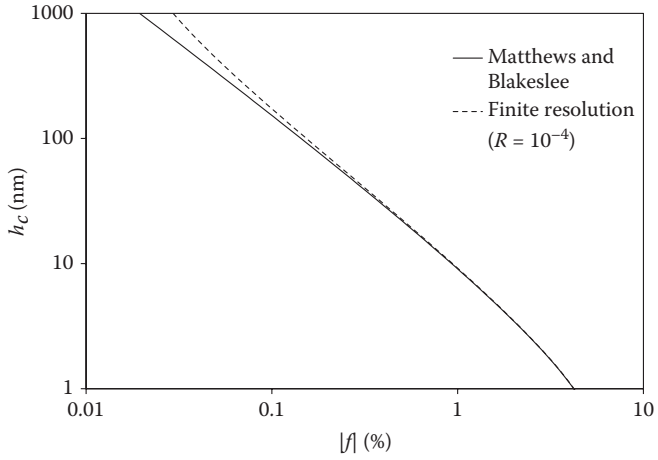


FIGURE 7.40
CLT vs. the lattice mismatch strain *f*. The solid line shows the Matthews and Blakeslee[109] CLT h_c; the dashed line shows the apparent CLT h_c^*, which would be measured using an experimental technique with a resolution $R = 10^{-4}$.

$$|f| - \frac{R}{Q} = \frac{b\left(1 - \nu\cos^2\alpha\right)\left[\ln\left(h_c^*/b\right) + 1\right]}{8\pi h_c^*\left(1-\nu\right)\cos\lambda}. \tag{7.111}$$

Fritz simplified this expression for (001) heteroepitaxy of diamond or zinc blende semiconductors to

$$|f| - \frac{R}{Q} = \frac{0.22}{h_c^*}\left[\ln\left(\frac{h_c^*}{4}\right) + 1\right], \tag{7.112}$$

where the apparent CLT h_c^* is in Angstroms.

Anomalously large CLT values have been determined using strain-based methods, for InGaAs/GaAs(001) by Orders and Usher[99] and Laidig et al.,[110] and for SiGe/Si(001) by People and Bean.[103] Fritz showed that these results can be (approximately) reconciled with the Matthews and Blakeslee[109] CLT if it is assumed that $R/Q = 7.5 \times 10^{-3}$. Figure 7.41 shows the previously mentioned experimental results, along with the apparent CLT curve calculated assuming $R/Q = 7.5 \times 10^{-3}$.

On the other hand, the anomalously large CLTs reported in XRD studies could be due to errors introduced by the crystallographic tilting of the epitaxial layers.[106] This potential source of error in x-ray measurements, now understood, can easily be removed by averaging peak separations from rocking curves taken at opposing azimuths.[107,108]

Therefore, the apparent CLT determined by any experimental method will generally be larger than the actual value in the presence of initially sluggish strain relaxation. The extent of this effect is determined by the ratio R/Q, where R is the experimental resolution and Q is the ratio of the misfit dislocation density to the density expected in equilibrium. The physics of the parameter Q is poorly understood, however. And whereas Fritz assumed Q to be constant, it is actually a function of time and layer thickness. Considerable work

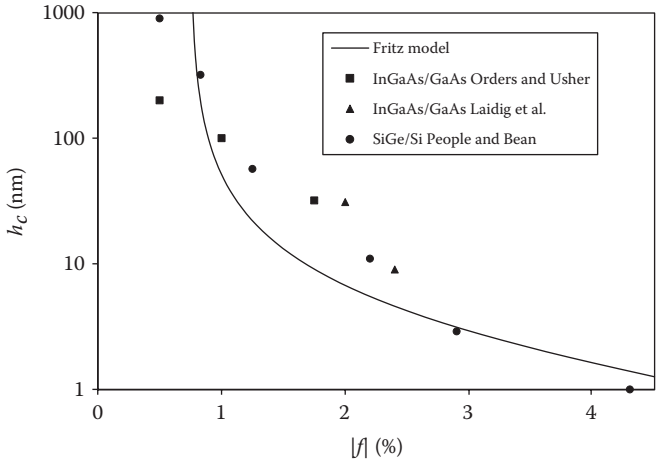


FIGURE 7.41
Apparent CLT vs. mismatch curve. The curve was calculated using the model of Fritz, with $R/Q = 7.5 \times 10^{-3}$. (The squares are data from P. J. Orders and B. F. Usher, *Appl. Phys. Lett.*, 50, 980 [1987]; the triangles are data from W. D. Laidig et al., *J. Vac. Sci. Technol. B*, 2, 181 [1984]; and the circles are data from R. People and J. C. Bean, *Appl. Phys. Lett.*, 47, 322 [1985]; *Appl. Phys. Lett.*, 49, 229 [1986]. Adapted from I. J. Fritz, *Appl. Phys. Lett.*, 51, 1080 [1987]. With permission. Copyright 1987, American Institute of Physics.)

remains before these issues can be satisfactorily resolved since experimental results for heteroepitaxial materials show that the lattice relaxation is not a fixed fraction of the equilibrium relaxation. In the general case, if we remove the assumption that δ/δ_{eq} is constant, the relationship between the resolution R and the measured CLT is

$$R = f - \varepsilon_{\parallel}\left(h_c^*\right). \tag{7.113}$$

Generally, for a particular set of growth conditions, the use of a kinetic model for lattice relaxation by dislocation flow in semiconductor heterostructures may enable a better understanding of the experimental resolution associated with the apparent CLT. Kujofsa et al.[111] investigated strain relaxation versus thickness and found the apparent CLT as a function of experimental resolution for a given growth temperature. The paper concluded that the growth temperature strain is important in determining the rate of relaxation, and this relaxation is not offset during cooldown. Furthermore, ZnSe layers grown at higher temperatures exhibit a much more rapid decrease in the residual strain, with thickness compared with those grown at lower temperatures. Because of this, measurements of the same apparent critical thickness will require much better experimental resolution with a growth temperature of 300°C, compared with the case of 595°C.

7.9.2 X-Ray Diffraction Determination of the CLT

The critical thickness has been determined using x-ray rocking curves based on strain[99,112,113] and also the rocking curve FWHM.[112] In either case, a series of heteroepitaxial samples is produced with a range of layer thicknesses. A rocking curve is measured for each sample, and the CLT is deduced from the strain versus thickness or rocking curve width versus thickness characteristic.

7.9.2.1 X-Ray Strain Method for CLT Determination

In the strain method, values of the in-plane elastic strain are determined from the separation of the substrate and epitaxial layer diffraction peaks. It is most convenient to use a symmetric x-ray reflection for this purpose. For example, the 004 reflection is typically used for (001) heteroepitaxy of diamond or zinc blende semiconductors, and the 0002 reflection is often used for (0001) wurtzite semiconductors.

The resolution of the XRD strain method is typically 10^{-5} to 10^{-4}, but depends on a number of factors, as will be shown in detail below. This level of resolution allows accurate determination of the CLT, so long as the difference in Bragg angles between the epitaxial layer and substrate is determined correctly. Large errors can be introduced by the crystallographic tilting between the epitaxial layer and substrate if this is not properly accounted for. This is because the rocking curve peak separation $\Delta\theta$ is given by

$$\Delta\theta = \Delta\theta_B + \Delta\phi\cos(\psi - \psi_0), \tag{7.114}$$

where:

$\Delta\theta_B = \theta_{Be} - \theta_{Bs}$ is the difference in Bragg angles between the epitaxial layer and substrate
ψ is the azimuthal angle for the incident x-ray beam
$\Delta\phi$ is the crystallographic tilting between the epitaxial layer and the substrate
ψ_0 specifies the direction of the epitaxial layer tilt

In early work, x-ray rocking curve results were analyzed with the assumption that $\Delta\theta \approx \Delta\theta_B$. However, this is only correct in the absence of epitaxial layer tilting, and leads to large errors in most cases. It is therefore necessary to find the Bragg angle difference by measuring the rocking curve at two opposing azimuths, and then taking the average:

$$\Delta\theta_B = \frac{\Delta\theta(\psi = 0^\circ) + \Delta\theta(\psi = 180^\circ)}{2}. \tag{7.115}$$

It should be noted that Equation 7.115 applies only in the case of a symmetric reflection. If an asymmetric reflection is used, an additional tilting of the diffracting planes is introduced by tetragonal distortion of the epitaxial layer. This term does not zero out when the rocking curve peak separations are averaged from opposing azimuths, as is detailed in Section 7.8.2.

The out-of-plane lattice constant c for the epitaxial layer can be determined using the Bragg law,

$$2d\sin\theta_B = n\lambda, \tag{7.116}$$

where:

d is the interplanar spacing
θ_B is the Bragg angle
n is the order of the reflection
λ is the x-ray wavelength

Often, Cu kα_1 radiation is used, with $\lambda = 0.1540594$ nm. For the 004 reflection from a zinc blende semiconductor, $c = d$ and $n = 4$. For the 0002 reflection from a wurtzite semiconductor, $c = d$ and $n = 2$. If it is assumed that the substrate is unstrained, and its lattice constant is known, then the out-of-plane lattice constant for the epitaxial layer may be determined by applying the Bragg law to each.

The out-of-plane strain in the epitaxial layer is then

$$\varepsilon_{\perp} = \frac{c - a_e}{a_e}, \tag{7.117}$$

where a_e is the relaxed lattice constant of the epitaxial layer. Assuming isotropic elasticity and tetragonal distortion of the epitaxial layer, the in-plane strain can be found by

$$\varepsilon_{\parallel} = -\left(\frac{2\nu}{1-\nu}\right)\varepsilon_{\perp}, \tag{7.118}$$

where ν is the Poisson ratio.

The sensitivity of the x-ray strain method is limited by the uncertainty in the peak separation $\Delta\theta_B$ between the diffraction peaks of the epitaxial layer and the substrate. In other words, the minimum detectable peak shift of the epitaxial layer rocking curve leads to an uncertainty $\Delta\varepsilon_{\perp}$ in the calculated out-of-plane strain $\varepsilon_{\perp}$ and a corresponding uncertainty $\Delta\varepsilon_{\parallel}$ in the in-plane strain $\varepsilon_{\parallel}$.

The resolution of the strain method may be analyzed as follows. By differentiating the Bragg law with respect to θ_B, we obtain

$$\frac{\Delta d}{d} = -\Delta\theta \cot\theta_B. \tag{7.119}$$

Then, if the strained layer is tetragonally distorted, then with a symmetric reflection we obtain

$$\Delta\varepsilon_{\perp} = \frac{\Delta c}{a_e} \approx \frac{\Delta c}{c} = \frac{\Delta d}{d} = -\Delta\theta \cot\theta_B. \tag{7.120}$$

Zhang et al.[113] found that the uncertainty $\Delta\theta$ in the peak position for the epitaxial layer was directly proportional to the epitaxial layer rocking curve FWHM, β. This is expected to hold in the general case, to the extent that the rocking curves have approximately the same shape, because a narrower rocking curve has a more rapidly changing first derivative. For a pseudomorphic heteroepitaxial layer, in which there are no significant sources of rocking curve broadening other than the finite layer thickness, the uncertainty in the Bragg angle is $r\beta_0$, where β_0 is the rocking curve width for a perfect crystal and r is a unitless constant of proportionality. Zhang et al. determined the constant of proportionality to be $r = 1/25$ in their experiments. However, this value may vary depending on the experimental conditions (effective number of x-ray counts observed) and diffractometer being used, as well as the profile shape.

For a pseudomorphic heteroepitaxial layer having a thickness h_c, the uncertainty in the epitaxial layer Bragg angle is

$$\Delta\theta_B \approx \pm r\beta_0(h_c), \tag{7.121}$$

where $\beta_0(h_c)$ is given by the Scherrer formula.[114,115] For this development, it is necessary to use the Scherrer equation in order to obtain an analytical result. For its application, however, either the Scherrer equation or the results of dynamical simulations may be used.

According to the Scherrer formula, the FWHM for a perfect pseudomorphic epitaxial thin layer is given by

$$\beta_0 \approx \frac{0.9\lambda}{h\cos\theta_B} \tag{7.122}$$

so that

$$\Delta\theta_B \approx \pm r\beta_0(h_c) = \frac{0.9r\lambda}{h_c\cos\theta_B}. \tag{7.123}$$

Substituting Equation 7.123 into Equation 7.120, we obtain the uncertainty in the calculated out-of-plane strain as

$$\Delta\varepsilon_\perp = -\Delta\theta_B = \mp\frac{0.9r\lambda}{h_c\sin\theta_B}. \tag{7.124}$$

The minimum lattice relaxation that may be detected using the x-ray strain method is

$$R(\text{strain}) = |\Delta\varepsilon_\parallel| = \left(\frac{2\nu}{1-\nu}\right)|\Delta\varepsilon_\perp| = \left(\frac{2\nu}{1-\nu}\right)\left(\frac{0.9r\lambda}{h_c\sin\theta_B}\right). \tag{7.125}$$

In the case of (001) heteroepitaxy of a diamond or zinc blende semiconductor using the 004 reflection and Cu kα_1 radiation ($\lambda = 0.1540594$ nm), assuming $r = 1/25$, $\nu = 1/3$, $h_c = 200$ nm, and $\theta_B = 33°$, the minimum detectable lattice relaxation is $R \approx 5 \times 10^{-5}$.

Orders and Usher[99] applied the XRD strain method to find CLTs in MBE-grown $In_xGa_{1-x}As$/GaAs(001) layers with various compositions. In this material system, the room temperature mismatch strain is approximately $f \approx x7.2\%$. Using 004 double-crystal rocking curves measured at a single azimuth, they estimated the strains in layers having various thicknesses for several particular compositions. Some representative rocking curves are shown in Figure 7.42. In this way, they determined the CLTs to be 200, 100, and 30 nm, for $x = 0.07$, 0.14, and 0.25, respectively. In their study, the composition x for each series of samples was determined by the rocking curve peak separation for thin layers. The strains were also calculated from the rocking curve peak separations. Both determinations would be subject to errors introduced by the crystallographic tilting of the epitaxial layers, however.

Petruzzello et al.[90] applied the XRD strain method to determine the critical thickness in MBE-grown ZnSe/GaAs(001). (For this heteroepitaxial system, the lattice mismatch strain is $f = -0.27\%$ at room temperature and the Matthews and Blakeslee[109] CLT is $h_c = 47$ nm.) For this purpose, they used a precision biaxial diffractometer[116] with absolute angle encoders and thereby avoided problems of crystallographic tilting between the ZnSe and GaAs. The strain versus thickness characteristic they measured in this way (Figure 7.43) was corroborated by TEM, PL, and powder diffractometer measurements. They found that the 87 nm layer was pseudomorphic, but the 180 nm layer was partially relaxed. From this observation, it is clear that the CLT is between 87 and 180 nm. The value of the CLT was estimated to be 150 nm, but no samples close to this thickness were grown in order to refine this estimate.

Zhang et al.[113] also used the XRD strain method to determine the CLT in MOVPE-grown ZnSe/GaAs(001). They used a Bartels high-resolution diffractometer and corrected for the

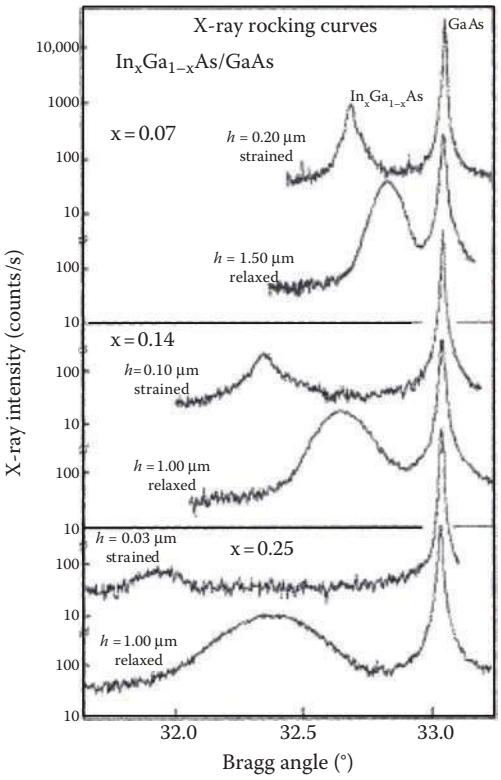


FIGURE 7.42
004 X-ray rocking curves from $In_xGa_{1-x}As/GaAs(001)$ heterostructures. (Adapted from P. J. Orders and B. F. Usher, *Appl. Phys. Lett.*, 50, 980 [1987]. With permission. Copyright 1987, American Institute of Physics.)

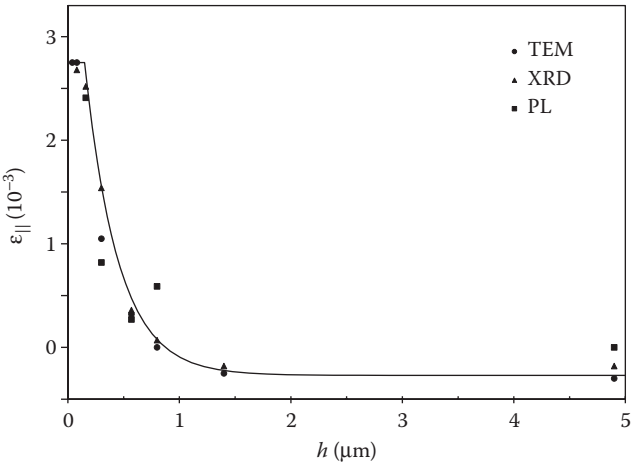


FIGURE 7.43
Room-temperature in-plane elastic strain in ZnSe/GaAs(001) measured by XRD, TEM, and PL. (Adapted from J. Petruzzello et al., *J. Appl. Phys.*, 63, 2299 [1988]. With permission. Copyright 1988, American Institute of Physics.)

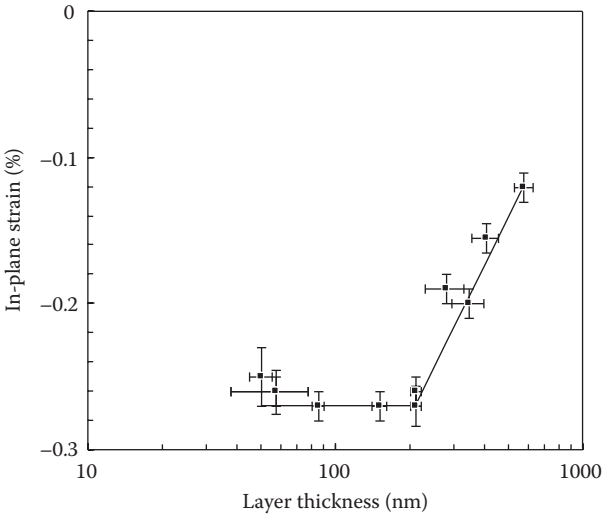


FIGURE 7.44
X-ray strain method for the determination of the CLT, applied to ZnSe/GaAs(001). Here, the room temperature in-plane strain has been plotted as a function of the layer thickness, for ZnSe/GaAs(001). The solid lines are guides to the eye. (Reprinted from X. G. Zhang et al., *J. Electron. Mater.*, 28, 553 [1999]. With permission. Copyright 1999, Minerals, Metals and Materials Society.)

epitaxial layer tilt. Figure 7.44 shows the in-plane strain determined at room temperature as a function of the layer thickness, for ZnSe/GaAs(001). The solid lines are guides to the eye. Here, the apparent CLT is approximately 210 nm.

Kim et al.[117] applied the XRD strain method to determine the CLT for GaN on AlN, in GaN/AlN/α-Al_2O_3 (0001) structures. (The room temperature in-plane lattice mismatch strain for this combination is $f = -2.6\%$.) In this study, MBE was used to produce GaN layers of varying thickness (50 Å to 1 μm) on 3.2 nm thick AlN buffer layers on *c*-face sapphire substrates. The thickest GaN layer was measured using selective etching and a mechanical profilometer, with 5% precision. The thicknesses of the remaining layers were estimated assuming the thickness was linear in the growth time. XRD measurements were made using synchrotron radiation at the National Synchrotron Light Source (NSLS). Multiple reflections were used and least-squares fitting was employed to improve the reliability of the results. They found the in-plane lattice constant fit the theoretical expression for equilibrium films,

$$a = a_s + \frac{h_c}{h}(a_s - a_e), \tag{7.126}$$

with a CLT of $h_c = 2.9 \pm 0.4$ nm. This value is in agreement with the prediction of the Matthews and Blakeslee[109] model ($h_c = 2.8$ nm, assuming $\nu = 0.38$, $b = 3.084$ Å, $\lambda = 30°$, and $\alpha = 90°$).

Their data are shown in Figure 7.45, along with the curve calculated using Equation 7.126 and the measured in-plane lattice constant for the AlN buffer layer ($a_s = 3.084$ Å).

7.9.2.2 X-Ray FWHM Method for CLT Determination

With the FWHM method, the determination of h_c is based on the observation of the rocking curve broadening by dislocations that are introduced during the relaxation process.

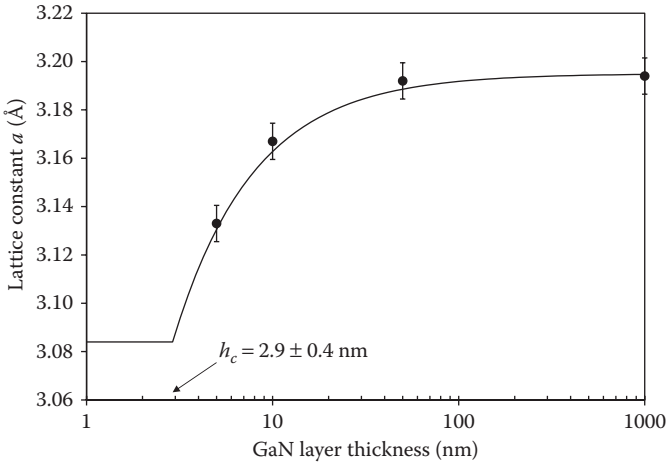


FIGURE 7.45
Room temperature in-plane lattice constant *a* vs. thickness for GaN on AlN, in GaN/AlN/α-Al_2O_3(0001) structures grown by MBE. (Adapted from C. Kim et al., *Appl. Phys. Lett.*, 69, 2358 [1996]. With permission. Copyright 1996, American Institute of Physics.)

It involves a comparison of the experimental rocking curve widths with the expected widths for perfect crystal layers. The perfect crystal rocking curve widths should usually be determined using dynamical simulations.[26,118,119] The Scherrer equation may also be used as long as the layer thickness is much less than the extinction depth for the x-rays.

For a dislocation-free pseudomorphic epitaxial layer, the width of the x-ray rocking curve decreases monotonically with the layer thickness. At the onset of lattice relaxation, misfit and threading dislocations are typically introduced together by the heterogeneous nucleation of dislocation half loops or dislocation multiplication processes. The dislocations cause significant broadening of the x-ray rocking curve when the epitaxial layer thickness is beyond the CLT for dislocation multiplication. This phenomenon is well known and is the basis for the x-ray determination of threading dislocation densities in heteroepitaxial layers.[120–122]

In the following analysis, it was assumed that the dislocation broadening is due to half loops and bent-over substrate dislocations, which participate in the strain relaxation process. In either case, the dislocations have threading and misfit segments, both of which can broaden the rocking curves. Kaganer et al.[173] presented an analysis of the rocking curve broadening by the misfit segments for some particular configurations. However, it is expected that the broadening will be dominated by the threading segments in most cases.

Commonly, the threading dislocation densities in heteroepitaxial diamond and zinc blende layers are determined from the broadening of the 004 rocking curve. This involves fitting the 004 rocking curve with a Gaussian peak, to determine the FHWM, and then deconvoluting the contribution due to the intrinsic rocking curve width for a perfect crystal.[121] Hence, the broadening due to the dislocations is given by

$$\beta_d = \sqrt{\beta^2 - \beta_0^2}, \tag{7.127}$$

where:

β is the experimentally measured FWHM
β_0 is the intrinsic rocking curve FHWM

By differentiating this equation with respect to β_d, we obtain the uncertainty in the dislocation broadening, which is

$$\Delta\beta_d = \left(\frac{2\beta}{\beta_d}\right)\Delta\beta. \tag{7.128}$$

The uncertainty in the rocking curve width, $\Delta\beta$, was found by Zhang et al. to be proportional to the measured rocking curve width, with the same constant of proportionality as in Equation 7.121. For a layer at the CLT,

$$\Delta\beta \approx \pm r\beta_0(h_c). \tag{7.129}$$

It was assumed that the minimum detectable threading dislocation broadening $\Delta\beta_{d\min}$ corresponds to the uncertainty in the broadening $\Delta\beta_d$. Then,

$$\beta_{d\min} = \Delta\beta_d = \left[\frac{2\sqrt{\beta_0^2 + \beta_{d\min}^2}}{\beta_{d\min}}\right]\Delta\beta. \tag{7.130}$$

Solving,

$$\beta_{d\min} = \frac{\Delta\beta\left[1+\sqrt{1+4(\beta_0/\Delta\beta)^2}\right]^{1/2}}{\sqrt{2}} = \frac{\Delta\beta\left[1+\sqrt{1+(2/r)^2}\right]^{1/2}}{\sqrt{2}}. \tag{7.131}$$

Assuming $r \ll 1$, the minimum detectable threading dislocation broadening for a layer with thickness h_c is

$$\beta_{d\min} \approx \frac{\Delta\beta}{\sqrt{r}} = \beta_0(h_c)\sqrt{r}. \tag{7.132}$$

If we apply the Scherrer equation, we have

$$\beta_{d\min} \approx \frac{0.92\sqrt{r}}{h_c\cos\theta_B}. \tag{7.133}$$

From the minimum detectable dislocation broadening $\beta_{d\min}$, the corresponding dislocation density $D_{\min}$ may be estimated as follows. If $\tan^2\theta_B < 2$, then the rocking curve broadening is dominated by the angular mosaic spread rather than the d-spacing mosaic spread. For this situation, β_d is approximately independent of the Bragg angle, and

$$D_{\min} = \frac{\beta_{d\min}^2}{4.36b^2}, \tag{7.134}$$

where b is the length of the Burgers vector for the threading dislocations.

To compare the FWHM and strain methods in terms of sensitivity, it is necessary to relate $D_{\min}$ from the former method to $\Delta\varepsilon_{\|}$ obtained from the latter. The relationship between the misfit dislocation density ρ_{MD} and the in-plane strain $\varepsilon_{\|}$ is

$$\varepsilon_{\parallel} = f - \rho_{MD}\, b \sin\alpha \sin\phi\,, \tag{7.135}$$

where:

- b is the length of the Burgers vector
- α is the angle between the Burgers vector and line vector
- ϕ is the angle between the interface and the normal to the slip plane

If it is assumed that there are two orthogonal misfit dislocation arrays with the same average misfit dislocation length L, and that each misfit dislocation has n threading dislocations associated with it, then

$$\rho_{MD} = \frac{DL}{2n}, \tag{7.136}$$

where D is the threading dislocation density. The value of n can be 0 (both ends of the misfit dislocation terminate at the sides or bottom of the sample), 1 (corresponding to the bending over of a substrate threading dislocation), or 2 (corresponding to a half loop introduced at the surface of the epitaxial layer). The average length of the misfit dislocations L may be estimated if it is assumed that they lie along ⟨110⟩ directions in the (001) interface and all have equal lengths. Then,

$$L = \sqrt{\frac{2n}{D}}. \tag{7.137}$$

Combining the previous three equations, we can find the minimum detectable strain relaxation associated with the FWHM method. It is

$$R(FWHM) = \Delta\varepsilon_{\parallel} = \frac{0.15\lambda\sqrt{r}}{h_c \cos\theta_B \sqrt{n}}. \tag{7.138}$$

In the case of (001) heteroepitaxy of a diamond or zinc blende semiconductor using the 004 reflection and Cu kα_1 radiation ($\lambda = 1.540594$Å), assuming $n = 2$, $r = 1/25$, $h_c = 200$ nm, and $\theta_B = 33°$, the minimum detectable lattice relaxation is about 1.5×10^{-5}. For this situation, the FWHM method turns out to be about three times more sensitive than the strain method, when using the same x-ray diffractometer.

For the case of their diffraction conditions, Zhang et al.[113] showed that the sensitivities of the two x-ray methods for CLT determination may be compared using the ratio

$$P = \frac{R(\text{strain})}{R(\text{FWHM})} = \frac{12\nu\sqrt{rn}\cot\theta_B}{1-\nu}. \tag{7.139}$$

If $P > 1$, then the FWHM will be more sensitive, but if $P < 1$, the strain method will be more sensitive. Typically, $P \approx 3$ and the FWHM method is preferred. Based on this, we can expect CLTs determined by the FWHM method to be smaller than those determined using the strain method, if the same diffractometer and experimental conditions are used.

More generally, Kujofsa and Ayers[45] showed that the sensitivity ratio depends on the observed number of x-ray counts, as well as the diffraction profile shape. In the general case, the sensitivity ratio is given by

$$P^G = \frac{R(\text{strain})}{R(\text{FWHM})} = \frac{\nu\sqrt{4.36x}\cot(\theta_B)}{(1-\nu)\cos(\alpha)\cos(\phi)}\sqrt{-N_{\text{Eff}}\ln(2)\ln\left(1-\frac{1}{\sqrt{N_{\text{Eff}}}}\right)} \tag{7.140}$$

and

$$P^L = \frac{R(\text{strain})}{R(\text{FWHM})} = \frac{\nu\sqrt{4.36x}\cot(\theta_B)}{(1-\nu)\cos(\alpha)\cos(\phi)}\sqrt{\frac{N_{\text{Eff}}}{N_{\text{Eff}}-1}}, \tag{7.141}$$

for the Gaussian and Lorentzian peak shapes, respectively.

Zhang et al.[113] compared the x-ray FWHM and strain methods for the determination of the CLT for ZnSe/GaAs(001) grown by photoassisted MOVPE. (For this heteroepitaxial system, the lattice mismatch strain is $f = -0.27\%$ at room temperature and the Matthews and Blakeslee[109] CLT is $h_c = 47$ nm.) Figure 7.46 shows the experimentally determined 004 rocking curve FWHM versus the ZnSe thickness (squares with error bars). Also shown are the calculated FWHM values for perfect crystal layers of ZnSe on GaAs(001), based on the Scherrer formula. It can be seen that the thin layers exhibit rocking curve widths that are the same as perfect crystal values, within the experimental errors. Significant dislocation broadening was observed for layers thicker than 135 nm. They considered the CLT to be the thickness at which the slope of the experimental characteristic (FWHM vs. thickness) was zero, giving a CLT of 140 nm. This result is comparable with those reported by Kamata and Mitsuhashi.[123] (~150 nm) and Petruzzello et al. (150 nm), and significantly less than that reported by Reisinger et al. (225 ± 5 nm). It was noted that this thickness corresponded to the onset of significant dislocation

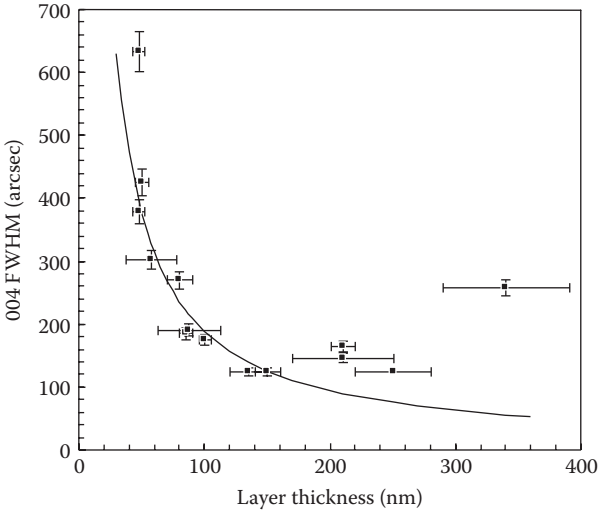


FIGURE 7.46
X-ray FWHM method for the determination of the CLT, applied to ZnSe/GaAs(001). Here, the 004 rocking curve FWHM has been plotted as a function of the layer thickness. The squares with error bars represent experimental data. The solid line was calculated using the Scherrer formula, for perfect crystals of ZnSe. (Reprinted from X. G. Zhang et al., *J. Electron. Mater.*, 28, 553 [1999]. With permission. Copyright 1999, Minerals, Metals and Materials Society.)

multiplication, rather than the CLT for the initiation of the glide of substrate dislocations. On the other hand, Zhang et al.'s result is significantly larger than the value of 97 nm reported by O'Donnell and coworkers.

The layer thickness for significant (observable) dislocation multiplication varies somewhat with growth conditions. Nonetheless, for a given heteroepitaxial system, a smaller value of CLT suggests better experimental sensitivity. Therefore, the x-ray FWHM method appears to have sensitivity superior to that of the x-ray strain method or the PL method. Only synchrotron XRT appears to be more sensitive than the x-ray FWHM method. However, all of the experimentally measured values of the CLT are significantly greater than the value predicted by Matthews and Blakeslee[109] (h_c = 47 nm for ZnSe/GaAs(001)) for the initial onset of lattice relaxation by the bending over of threading dislocations.

Zhang et al. also determined the CLT from the same set of samples using the x-ray strain method. The apparent CLT found in this way was 210 nm. However, they determined the XRD FWHM method ($R = 1.5 \times 10^{-5}$) to have better resolution than the XRD strain method ($R = 5 \times 10^{-5}$). The value of 140 nm was therefore considered more reliable.

7.9.2.3 X-Ray Topography Determination of CLT

XRT involves the capture of a photographic image from a diffracting crystal, set at the Bragg condition, and in reflection mode.[124] A collimated, monochromatic x-ray beam is incident on the crystal at the Bragg angle, diffracted, and collected by a photographic emulsion. A planar view is normally obtained unless the sample is specially prepared. Strained regions surrounding a dislocation diffract less intensity, and misfit dislocations are revealed in the image as dark lines. XRT can therefore reveal the creation of a single misfit dislocation in the sample.

The absolute resolution of this technique depends on the area examined. For example, if the heteroepitaxial layer is imaged over an area of $D \times D$, and a single misfit dislocation can be detected in this area, then the resolution is

$$R = \frac{b \sin \alpha \sin \phi}{D}, \tag{7.142}$$

where:

- b is the length of the Burgers vector
- α is the angle between the Burgers vector and line vector
- ϕ is the angle between the interface and the normal to the slip plane

For (001) heteroepitaxy of zinc blende semiconductors, an image size of 25 × 25 μm results in a resolution of $R = 8 \times 10^{-6}$.

O'Donnell et al.[125] used XRT to determine the CLT for MBE-grown ZnSe/GaAs(001). (For this material system, the room temperature mismatch strain is −0.27% and the Matthews and Blakeslee[109] CLT is 47 nm.) 044 topographs were obtained using synchrotron radiation at a wavelength of 0.148 nm. Samples of various thicknesses were prepared, and the layer thicknesses were calibrated by comparing the measured x-ray rocking curves with dynamical simulations. In a 95 nm thick heteroepitaxial layer, no misfit dislocations were imaged. Misfit dislocations were found to be present in a 100 nm thick layer with a spacing of 14 μm. Based on these observations, the CLT was estimated to be 97 nm.

7.9.2.4 TEM Determination of the CLT

TEM[90,91] can be used to determine the CLT, by the examination of samples having various epitaxial layer thicknesses. The samples may be prepared in either planar or cross-sectional view. Cross-sectional micrographs allow the viewing of individual misfit dislocations. It is also possible to determine the Burgers vectors for individual dislocations by using TEM micrographs obtained with various diffraction vectors for the electrons, using the $\mathbf{b} \cdot (\mathbf{g} \times \mathbf{1}) = 0$ criterion for the extinction of dislocation contrast, where $\boldsymbol{b}$ is the Burgers vector, $\boldsymbol{l}$ is the dislocation line vector, and $\mathbf{g}$ is the diffraction vector. This capability is a unique advantage of the TEM method. Disadvantages of the TEM method are the destructive nature of sample preparation and the potential for the introduction of artifact defects during this preparation.

Edirisinghe et al.[126] used TEM to establish the CLT in MBE-grown InGaAs/GaAs(111)B. They examined samples of different thicknesses and compositions in plan-view, bright-field micrographs obtained using the $(0\bar{2}2)$ diffraction vector. Figure 7.47 shows such a micrograph for a 150 nm thick $In_{0.15}Ga_{0.85}As$/GaAs(111)B sample. Misfit dislocations are clearly visible running along two of the ⟨110⟩ directions in the interface. However, the lattice relaxation is asymmetric due to the vicinal substrate, and so the misfit dislocations do not have equal densities in the two ⟨110⟩ directions. Also, misfit dislocations along the $[0\bar{1}1]$ direction are invisible with this diffraction vector.

Edirisinghe et al. found that, for $In_{0.25}Ga_{0.75}As$/GaAs(111)B, misfit dislocations were present in a 25 nm thick layer, but not in a 15 nm thick layer. This indicates that the CLT for this material system is between 15 and 25 nm. (The room temperature lattice mismatch strain is –1.8%.)

Petruzzello et al.[90] applied TEM to determine the critical thickness in MBE-grown ZnSe/GaAs(001). Here, cross-sectional and planar samples were studied, with layer thicknesses varying from 50 to 4900 nm. Samples with epitaxial layer thicknesses of 50 and 87 nm contained stacking faults on {111} planes bounded by Frank partial dislocations having Burgers vectors of the type $a/3\langle 111\rangle$. However, only the epitaxial layers with thicknesses of 180 nm or greater contained perfect misfit dislocations. From this observation, it is clear that the CLT is between 87 and 180 nm. The value of the CLT was estimated to be 150 nm, but no samples close to this thickness were grown in order to refine this estimate.

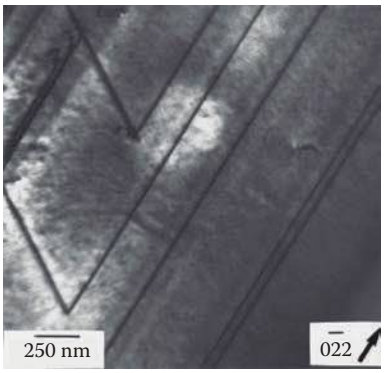


FIGURE 7.47
TEM plan-view, bright-field micrograph of a 150 nm thick layer of $In_{0.15}Ga_{0.85}As$/GaAs(111)B, with $\boldsymbol{g} = 0\bar{2}2$. Misfit dislocations are visible along two of the ⟨110⟩ directions in the interface. Misfit dislocations along the $[0\bar{1}1]$ direction are invisible with this diffraction vector. (Reprinted from S. P. Edirisinghe et al., *J. Appl. Phys.*, 82, 4870 [1997]. With permission. Copyright 1997, American Institute of Physics.)

Petruzzello et al. also used their TEM results to estimate the residual strains in the ZnSe/GaAs(001) heteroepitaxial layers as a function of the ZnSe thickness. This was done using the measured densities of 60° and edge dislocations in the samples. Whereas most of the misfit dislocations were of the 60° type, the fraction of Lomer-type edge dislocations increased with layer thickness. It was concluded that the edge dislocations originated from reactions between 60° misfit dislocations. The calculated residual strains were in agreement with high-resolution XRD measurements made using a high-resolution Biaxial diffractometer[115] and a conventional powder diffractometer, as well as PL results.

7.9.2.5 Electron Beam–Induced Current Determination of CLT

EBIC[101] allows the imaging of individual dislocations located in the vicinity of a p-n junction, and has been employed for CLT determination. Here, a p-n heterojunction is formed between the substrate and heteroepitaxial layer. Then, in a scanning electron microscope, an electron beam is scanned over the surface of the sample. The electron beam excites electron–hole pairs in the sample, which are separated by the built-in electric field of the zero-biased p-n heterojunction. The EBIC flowing in the substrate is measured and recorded at each position of the electron beam. Dislocations give rise to nonradiative recombination in the material, and a reduction in the EBIC. Therefore, if the magnitude of the current is used to make an image, misfit dislocations and other areas of short minority carrier lifetime show up as dark features on the EBIC image.

Kohama et al.[101] applied EBIC for the determination of CLTs in MBE-grown $Si_{1-x}Ge_x$/Si(001). The layers were grown on n-type Si substrates, and were unintentionally doped p-type, thus forming a p-n heterojunction at the interface. EBIC measurements were done in a SEM with 20 kV accelerating voltage and an electron beam current of 7×10^{-7} *A*.

In this study, EBIC and TEM were used to examine $Si_{1-x}Ge_x$/Si(001) samples having various thicknesses and compositions. The misfit dislocation lines imaged by the two techniques were correlated for a relaxed 300 nm thick layer of $Si_{0.8}Ge_{0.2}$/Si(001). Thinner layers were considered to be pseudomorphic if no misfit dislocations were found in the EBIC image. In this way, the greatest pseudomorphic thickness was plotted as a function of the lattice mismatch. For each composition, the smallest thickness in which there were misfit dislocations was also plotted. The actual CLT characteristic was considered to lie between the pseudomorphic and partially relaxed curves, as shown in Figure 7.48.

7.9.2.6 Photoluminescence Determination of the CLT

PL may be used to determine the CLT of a variety of structures, including single heteroepitaxial layers, strained single quantum wells (SSQWs), or strained layer superlattices. Several observable changes in the spectrum accompany lattice relaxation at the CLT, including a wavelength shift in the near-band-edge peak due to strain relaxation, the broadening and reduction in intensity of the near-band-edge peak, and in some cases, the appearance of deep-level emission. The CLT can be determined based on any of these phenomena or some combination. If the material being studied is to be used in optoelectronic devices, then the observed optical characteristics are immediately relevant.

If the peak emission energy is assumed to correspond to the bandgap in the material, a biaxial stress in the layer will induce a shift in the peak, which for a (001) zinc blende semiconductor is given by

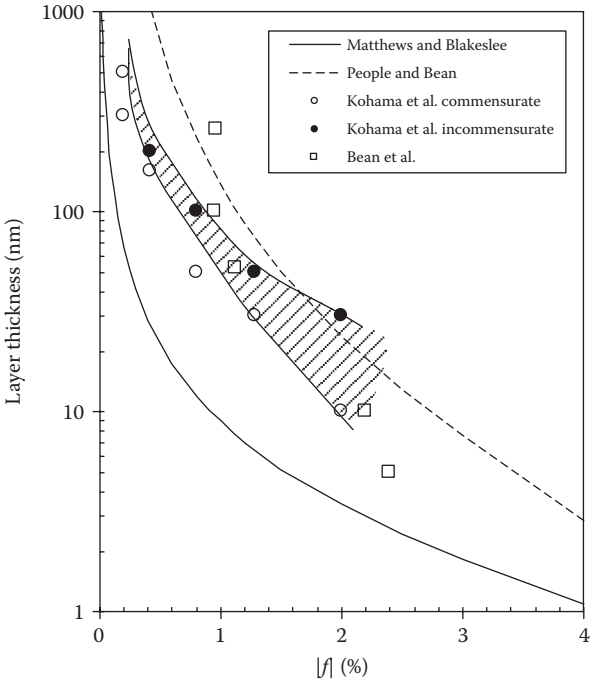


FIGURE 7.48
CLT vs. Ge mole fraction x in $Si_{1-x}Ge_x$/Si(001) as determined using EBIC. The solid dots represent the thinnest samples in which misfit dislocations were found. The open circles represent the thickest samples at each composition that were pseudomorphic (commensurate). Also shown are the experimental data of Bean et al.[132] and the predictions of the People and Bean[103] model and Matthews and Blakeslee[109] model. (Adapted from Y. Kohama et al., *Appl. Phys. Lett.*, 52, 380 [1988]. With permission. Copyright 1988, American Institute of Physics.)

$$\Delta E_g = -\varepsilon_{\parallel}\left[-2a\left(C_{11}-C_{12}\right)/C_{11} + b\left(C_{11}+2C_{12}\right)/C_{11}\right], \tag{7.143}$$

where:

$\varepsilon_{\parallel}$	is the in-plane strain
C_{11} and C_{12}	are the elastic constants
a and b	are the hydrostatic and shear deformation potentials, respectively

Usually, PL spectra are measured at cryogenic temperatures. In these cases, care must be taken to account for the temperature variation of the bandgap. Also, the strain at cryogenic temperatures will not be the same as the room temperature value due to the thermal strain. This can be accounted for if the thermal expansion characteristics are known for both substrate and deposit, but in some cases, this information is incomplete below 100K. Another complication arises in quantum wells, in which the (thickness-dependent) blue shift due to quantum confinement effects must be considered.

Parker et al.[127] used PL to determine the CLTs in $In_xGa_{1-x}N$ on GaN in $In_xGa_{1-x}N$/GaN/α-Al_2O_3 heterostructures grown by MOVPE. (In this combination, the room temperature lattice mismatch strain is $f = -x9.7\%$.) Film thicknesses were determined by TEM (thin layers) or optical measurements (thicker layers). Because of the negative (compressive) lattice mismatch strain, the near-band-edge emission was blue shifted in pseudomorphic layers. In partially relaxed layers, the blue shift decreased and deep-level emission

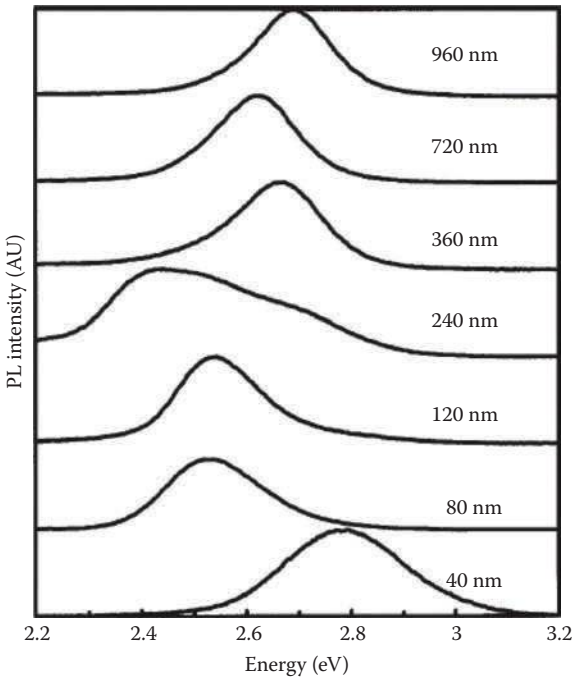


FIGURE 7.49
Room temperature PL spectra for $In_{0.15}Ga_{0.85}N/GaN$ layers with various InGaN thicknesses, in MOVPE-grown $In_xGa_{1-x}N/GaN/\alpha$-Al_2O_3 heterostructures. (Reprinted from C. A. Parker et al., *Appl. Phys. Lett.*, 75, 2776 [1999]. With permission. Copyright 1999, American Institute of Physics.)

eventually became dominant. Representative PL spectra are shown in Figure 7.49 for $In_{0.15}Ga_{0.85}N/GaN$ layers with various InGaN thicknesses. They considered the CLT to be that thickness for which the peak emission wavelength was the same as for a relaxed layer. Using this approach, they determined $h_c = 100$ nm for $x = 0.08$ and $h_c = 65$ nm for $x = 0.15$. These values are considerably larger than the predictions of the Matthews and Blakeslee model, which are 14 and 6.3 nm, assuming $\nu = 0.38$, $b = 3.1$Å, $\lambda = 30°$, and $\alpha = 90°$. However, a fairly limited number of layer thicknesses were studied. Also, the use of room temperature PL spectra resulted in strain resolution below that which can be obtained at cryogenic temperatures.

Reed et al.[94] used several techniques, including PL, to determine the CLTs for MOVPE-grown $GaN/In_xGa_{1-x}N/GaN$ quantum wells on sapphire (0001) substrates. Here, too, the wavelength shift of the peak intensity was used to deduce the CLT. The compositions were determined by XRD measurements on thick layers. They also estimated the CLTs based on the carrier mobility and quantum well conductivity, as found from Hall effect measurements. These results are shown in Figure 7.50 for indium compositions up to 16%.

7.9.2.7 PLM Determination of the CLT

PLM[57,61,95] is an imaging technique that allows the detection of individual misfit dislocations. It therefore allows detection of lattice relaxation at its very earliest stages, making PLM a highly sensitive method for determination of the CLT. The absolute resolution of this method depends on the area examined, in the same manner as for XRT.

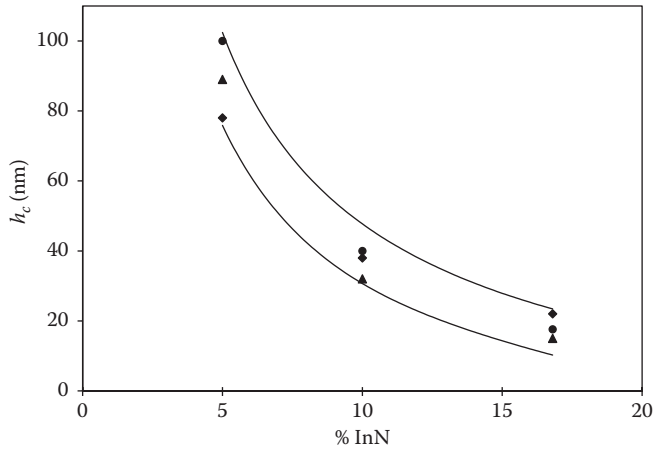


FIGURE 7.50
CLTs determined for $GaN/In_xGa_{1-x}N/GaN$ quantum wells on sapphire (0001) substrates, as a function of the quantum well indium composition. The CLTs were determined from PL spectra (diamonds), Hall effect quantum well carrier mobility (circles), and Hall effect quantum well conductivity (triangles). (Adapted from M. J. Reed et al., *Appl. Phys. Lett.*, *77*, 4121 [2000]. With permission. Copyright 2000, American Institute of Physics.)

Gourley et al.[95] applied PLM to determine the CLT in $In_{0.2}Ga_{0.8}As/GaAs(001)$ strained layers and superlattices. In this study, micro-PL images were obtained at 80 K for $In_{0.2}Ga_{0.8}As$ layer thicknesses varying from 50 to 600 Å. Figure 7.51 shows the PLM images for $GaAs/In_{0.2}Ga_{0.8}As/GaAs(001)$ SSQWs of various thicknesses. The SSQWs with thicknesses of 150 Å or less are free from dark-line defects (DLDs). However, SSQWs of 200 Å or greater show DLDs aligned with the ⟨110⟩ directions, indicating the presence of misfit dislocations at their interfaces.

From the PLM images, Gourley et al. determined and plotted the DLD density versus InGaAs thickness, for $GaAs/In_{0.2}Ga_{0.8}As/GaAs(001)$ SSQWs and also strained layer superlattices. These results are shown in Figure 7.52. For the SSQWs, the DLD density increases dramatically, and in approximately linear fashion, for thicknesses above 20 nm. By extrapolating the DLD density back to zero, a CLT of 19 nm was inferred. For the same samples, Gourley et al. used the x-ray strain method (Section 7.9.2) to determine the critical thickness and found a value between 40 and 60 nm. This is indicative of the excellent sensitivity afforded by the PLM technique.

7.9.2.8 RHEED Determination of the CLT

RHEED can be used for the *in situ* monitoring of lattice relaxation by a growing film in the MBE chamber. This contrasts with the previously described methods, which involve the examination of many samples having different thicknesses, outside of the growth chamber. Also, because RHEED can be used to monitor the layer thickness with atomic layer accuracy, the technique lends itself to high mismatch heteroepitaxial systems for which the CLT is of this order.

There are several features of RHEED characterization relevant to this purpose. First, the intensity of a particular RHEED spot is found to oscillate with time, and each period of oscillation corresponds to the growth of one monolayer. Second, a streaky RHEED pattern corresponds to an atomically smooth surface, whereas a spotty pattern indicates surface

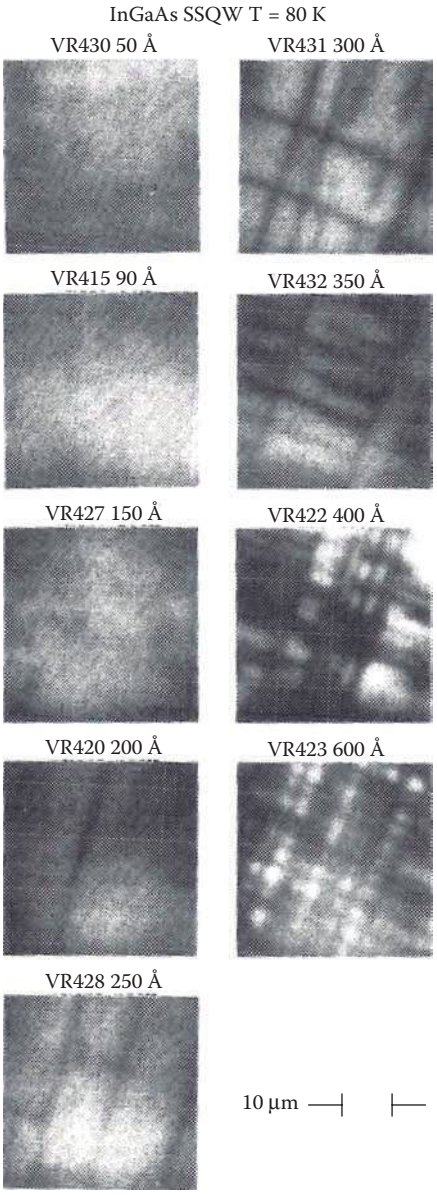


FIGURE 7.51
PLM images for GaAs/$In_{0.2}Ga_{0.8}As$/GaAs(001) SSQWs of various thicknesses, obtained at 80 K with 476.2 nm excitation with a continuous wave power density of 100 W cm^{-2}. (Reprinted from P. L. Gourley et al., *Appl. Phys. Lett.*, 52, 377 [1988]. With permission. Copyright 1988, American Institute of Physics.)

roughening. Third, the rapid damping of the RHEED intensity oscillations is associated with a change from two- to three-dimensional growth. Usually, the CLT for the two- to three-dimensional transition is assumed to be the same as the CLT for the onset of lattice relaxation by misfit dislocations, although these two will not necessarily be the same.

Elman et al.[128] used RHEED to investigate the CLT in $In_xGa_{1-x}As$/GaAs(001) grown by MBE, for various compositions and growth temperatures. Figure 7.53 shows their

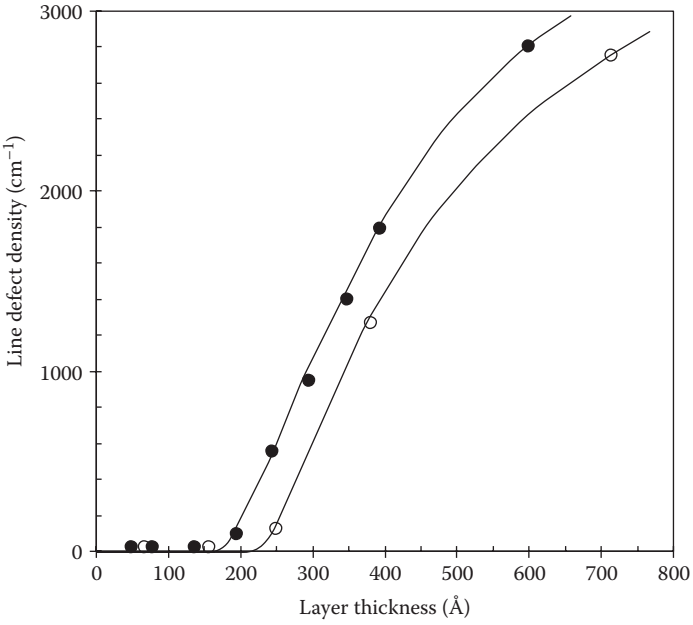


FIGURE 7.52
DLD densities as obtained from PLM images for GaAs/$In_{0.2}Ga_{0.8}As$/GaAs(001) SSQWs (solid points) and $In_{0.2}Ga_{0.8}As$/GaAs(001) strained layer superlattices (open circles) with various InGaAs layer thicknesses. (Reprinted from P. L. Gourley et al., *Appl. Phys. Lett.*, 52, 377 [1988]. With permission. Copyright 1988, American Institute of Physics.)

measured RHEED intensity oscillations for layers grown at 615°C with the compositions x = 0.24, 0.29, 0.32, 0.39, 0.42, and 0.50. From these characteristics, the CLTs were found in terms of monolayers for layers having these compositions and are shown in Figure 7.54, along with the predictions of the Matthews and Blakeslee model, and also PL results.

7.9.2.9 Scanning Tunneling Microscope Determination of the CLT

STM is often used to study the morphological aspects of relaxing mismatched layers.[129–131] However, under the appropriate conditions, this technique can also be used for the indirect observation of misfit dislocation formation during growth. STM involves the scanning of a fine metal tip over the surface of the sample, while the vertical position of the tip is adjusted to maintain a nearly constant tunneling current between the tip and the sample. While scanning, the vertical deflection of the tip is recorded as a function of position on the surface. In this way, a surface topographic map is produced for the sample, with atomic-scale resolution. Lattice distortions at the surface having subangstrom scale may be imaged as long as smooth layer-by-layer growth is achieved, and the CLT is on the order of a few atomic layers. Therefore, this method may only be applied to large-mismatch heteroepitaxial material systems. Also, the effective use of this method requires a purpose-built STM, which is integral to an ultra-high-vacuum MBE chamber, to allow examination of the deposit at various thicknesses. These requirements make the method applicable to only a few heteroepitaxial material systems.

Yamaguchi et al.[92] have used STM to determine the CLT in MBE-grown InAs/GaAs(111) A. (The room temperature mismatch strain for this combination is –7.2%.) In this material

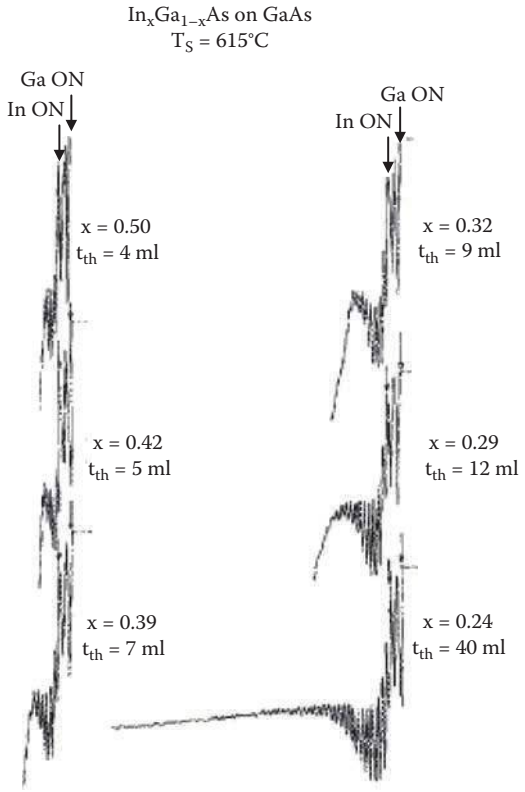


FIGURE 7.53
RHEED intensity oscillations for $In_xGa_{1-x}As$/GaAs(001) grown by MBE at a temperature of 615°C, with the compositions x = 0.24, 0.29, 0.32, 0.39, 0.42, and 0.50. (Adapted from B. Elman et al., *Appl. Phys. Lett.*, 55, 1659 [1989]. With permission. Copyright 1989, Institute of Physics.)

system, step-flow growth is maintained for all values of InAs coverage. Also, the InAs(111) A surface exhibits a (2 × 2) surface reconstruction with a corrugation of only 0.2 Å. Because of these characteristics, the misfit dislocations at the interface produce measurable distortions at the top surface.

Yamaguchi et al. found that the InAs grew heteroepitaxially in a step-flow mode, but with some nucleation of islands. At between 1 and 2 ML thickness, they observed the appearance of ~0.5 Å surface troughs at the coalesced boundaries of islands. Because the (2 × 2) reconstruction pattern was continuous across these troughs, they were interpreted to be the result of misfit dislocations at the heterointerface, rather than gaps between islands. Further, these misfit dislocations were found to form a trigonal network at the (111)A interface, with dislocation lines along the ⟨110⟩ directions. From this study, it was concluded that the CLT is between 1 and 2 ML (1 ML = 3 Å) for this material combination.

7.9.2.10 Rutherford Backscattering Determination of the CLT

The CLT may also be determined from Rutherford backscattering (RBS) experiments. Here, a He ion beam is incident on the sample, which can be adjusted in angle by a goniometer. Along certain low-index crystallographic directions such as ⟨100⟩ and ⟨110⟩, the

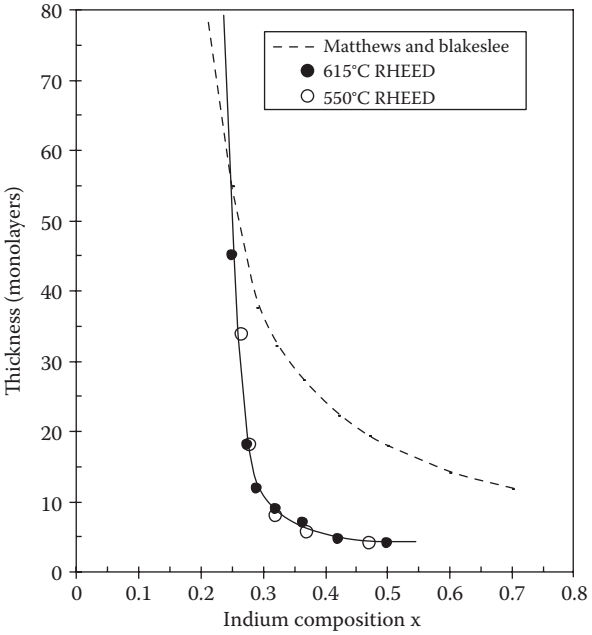


FIGURE 7.54
CLT (in monolayers) vs. composition x, for $In_xGa_{1-x}As/GaAs(001)$. (Adapted from B. Elman et al., *Appl. Phys. Lett.*, 55, 1659 [1989]. With permission. Copyright 1989, Institute of Physics.)

He ions channel deeply into the specimen, and only a small percentage of them will be backscattered to the detector. For other orientations (misaligned crystal), a large number of ions will be backscattered and detected. Therefore, if the number of backscattered ions per unit time is measured as a function of the angle of incidence for the He ion beam, there will be dips in the profile corresponding to the low-index directions in the crystal.

Figure 7.55 shows that the off-normal crystallographic directions in the heteroepitaxial layer will be shifted by the presence of strain. For a pseudomorphic, tetragonally distorted layer, as shown in Figure 7.55a, the ⟨110⟩ channeling direction for the heteroepitaxial layer will differ from that of the substrate. But for the relaxed layer of Figure 7.55b, the ⟨110⟩ directions will be parallel. Therefore, if the difference between the substrate and epitaxial layer channeling directions is measured for samples of different thicknesses, the CLT can be determined from the onset of observable strain relaxation.

For (001) heteroepitaxy of zinc blende or diamond crystals, if the ⟨110⟩ channeling directions of the epitaxial layer and substrate differ by $\Delta\phi_{tet}$, then the *tetragonal distortion* is given by

$$\frac{c-a}{a_e} = -\frac{\Delta\phi_{tet}}{\sin\phi\cos\phi}, \tag{7.144}$$

where:

c and a	are the out-of-plane and in-plane lattice constants for the epitaxial layer
a_e	is the relaxed lattice constant of the epitaxial layer
ϕ	is the angle between the ⟨110⟩ channeling direction and the normal to the surface

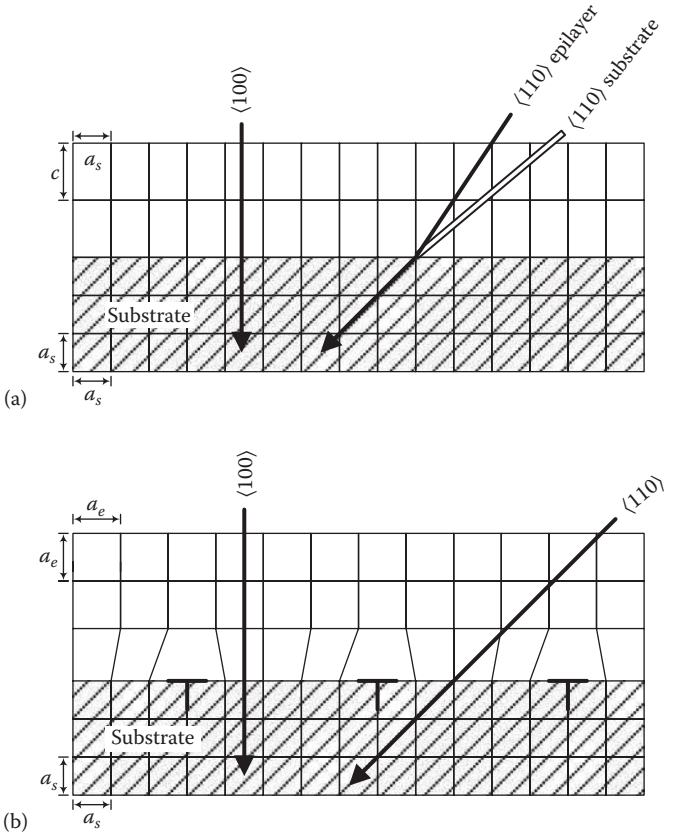


FIGURE 7.55
Schematic illustration of backscattering directions in mismatched heteroepitaxial structures. (a) For a pseudomorphic layer, the tetragonal distortion introduces a tilt of the ⟨110⟩ channeling directions relative to the substrate. (b) For a relaxed layer, this tilt is zero. (Adapted from J. C. Bean et al., *J. Vac. Sci. Technol. A*, 2, 436 [1984]. With permission. Copyright 1984, American Vacuum Society.)

For the ⟨110⟩ channeling direction, $\phi = 45°$.

Bean et al.[132] used RBS to determine the CLT in $Si_{1-x}Ge_x$/Si(001) and found good agreement with x-ray strain measurements.

7.10 Determination of the Crystal Orientation

In some cases, the crystal orientation of the heteroepitaxial layer may be very different from that of the substrate, or the two crystals may even have different crystal structures. In these situations, the measurement of three *hkl* XRD profiles (with three non-coplanar diffraction vectors) allows the determination of the crystal orientation, for a nonpolar heteroepitaxial layer such as Ge_xSi_{1-x}. For a polar semiconductor, a crystallographic etch may be used in conjunction with the x-ray measurements, for the unambiguous assignment of the heteroepitaxial relationship between the deposit and substrate. For example, in the

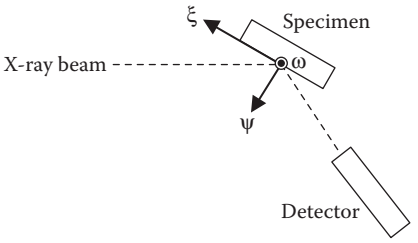


FIGURE 7.56
Geometry used in an HRXRD experiment. A rocking curve (ω scan) involves measuring the diffracted intensity as a function of the rocking angle ω. Prior to measurement of the rocking curve, the tilt ξ must be adjusted so that the diffraction vector is in the plane of the diffractometer. Rocking curves may be measured at different azimuthal positions ψ. However, the tilt must be readjusted for each new azimuth.

case of GaAs, molten KOH etching can be used to distinguish between the [110] and [$\bar{1}$10] directions.

More often, the epitaxial layer and substrate differ only slightly in orientation. Then, this misorientation may be found by measuring diffraction profiles at various azimuths. Figure 7.56 illustrates the specimen geometry used in an HRXRD experiment. ω, ξ, and ψ are the rocking, tilt, and azimuthal angles, respectively. Rocking curves are obtained by measuring the diffracted intensity as a function of the rocking angle ω. Rocking curves may be measured at different azimuths; for each setting of the azimuth ψ, the tilt ξ is adjusted so that the diffraction vector is in the plane of the diffractometer. At an azimuthal rotation ψ, the angular separation Δθ between the epitaxial layer rocking curve intensity peak and the substrate rocking curve peak will be

$$\Delta\theta = \Delta\theta_B + \Delta\phi\cos(\psi - \psi_0), \tag{7.145}$$

where:

$\Delta\theta_B$ is the difference in Bragg angles between the epitaxial layer and substrate
$\Delta\phi$ is the misorientation between the epitaxial layer and the substrate
ψ_0 specifies the direction of the misorientation

Therefore, if rocking curves are measured at three or more azimuths, a plot of Δω versus ψ allows the determination of the crystallographic orientation (Δϕ and ψ_0) for the epitaxial layer.

To illustrate the technique, Figure 7.57 shows the measured rocking curve peak separation as a function of azimuth, for a 4.7 μm thick layer of GaAs grown on Si(001) by MOVPE.[107] The filled squares represent the measured data. The curve is the best sinusoidal fit to the data, given by

$$\Delta\theta = 5270'' + 225''\cos(\psi - 225^\circ). \tag{7.146}$$

In this example, therefore, the GaAs layer is tilted by 225 in. (0.0625°) toward the [010]. Because the substrate was tilted by 2.25° toward the [0$\bar{1}$0], tilt of the [001] axis of the epitaxial layer with respect to that of the substrate is toward the surface normal.

Although rocking curves represent the most common approach for the characterization of crystal misorientation between an epitaxial layer and its substrate, Greenberg and Ladell[133] have demonstrated another x-ray technique, modulation of Renninger scan

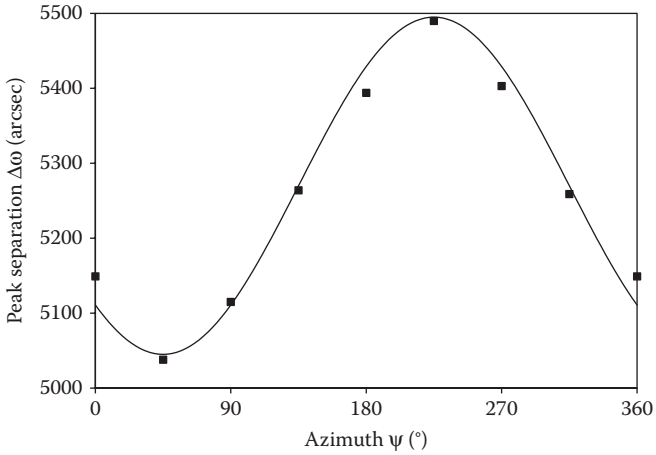


FIGURE 7.57
004 rocking curve peak separation Δθ vs. the azimuthal angle ψ, for 4.7 μm thick GaAs/Si(001) grown by MOVPE. The experimental data are shown with filled squares. Also plotted is the curve given by $\Delta\theta = 5270'' + 225'' \cos(\psi - 225°)$. (Adapted from S. K. Ghandhi and J. E. Ayers, *Appl. Phys. Lett.*, 53, 1204 [1988]. With permission. Copyright 1988, American Institute of Physics.)

intensity (MORSI), which enables the characterization of the orientation, strain, and structural perfection of a heteroepitaxial layer. This method involves measurement of the diffracted intensity from a symmetric *hkl* reflection while rotating the sample about its azimuth. This Renninger scan exhibits dips in intensity corresponding to multiple diffraction conditions, and an additional structure is introduced in the intensity scan by the interaction of the epitaxial layer and substrate.

7.11 Characterization of Defect Types and Densities

Defects such as dislocations, stacking faults, twins, and inversion domain boundaries (IDBs) may be investigated using a number of characterization techniques. TEM, crystallographic etching, and XRD are employed most often. Of these, TEM is the most powerful, as it can image individual dislocations or planar defects. It is possible with TEM to determine the Burgers vectors for dislocations, as well as their line vectors. However, the TEM can image only a small volume of the sample. In addition, the sample preparation is destructive and, in some cases, could alter the configuration of the defects being studied. Crystallographic etching has been used extensively to determine the densities of dislocations intersecting a crystal surface. The crystallographic etches reveal pits (or occasionally hillocks) at the points of emergence for threading dislocations. This technique can be used over a large sample area, but is not applicable for very high dislocation densities, for which the etch pits overlap. XRD is nondestructive and can be used to estimate the average dislocation density in the volume of the sample, from the dislocation broadening of the rocking curve widths. Other techniques, such as PLM, cathodoluminescence (CL), x-ray topography (XRT), EBIC, or scanning tunneling microscopy (STM), can be used to image individual dislocations. Their use has been covered in Section 7.9 and will not be repeated here.

7.11.1 Characterizing Defects by TEM

TEM is capable of imaging individual dislocations, similar to PLM, CL, XRT, and EBIC, but it also allows determination of the Burgers vector for a dislocation using the $\mathbf{g} \cdot \mathbf{b}$ analysis described in Section 7.4.2. This information can be used to understand the geometry of the slip systems in a semiconductor crystal, and the character of the misfit and threading dislocations.

Figure 7.58 illustrates the method of Burgers vector identification using TEM, from the work of Abrahams et al.[134] The sample investigated comprised a graded $GaAs_{1-x}P_x$ layer grown on GaAs(001) by vapor phase epitaxy. The two-beam, bright-field TEM micrographs were obtained for the identical area of a single sample, but with four different diffraction vectors. The misfit dislocation under the numeral 1 vanishes for the diffraction vector $\mathbf{g} = [02\bar{2}]$, but is visible for $\mathbf{g} = [0\bar{2}\bar{2}]$, $\mathbf{g} = [0\bar{4}0]$, and $\mathbf{g} = [00\bar{4}]$. Its Burgers vector must therefore be perpendicular to $[02\bar{2}]$, but not any of the other diffraction vectors. From this, it can be concluded that the Burgers vector is parallel to [011], and so $\mathbf{b} = (a/2)[011]$. The line of this dislocation is parallel to $[01\bar{1}]$, so it is has edge character. Following similar

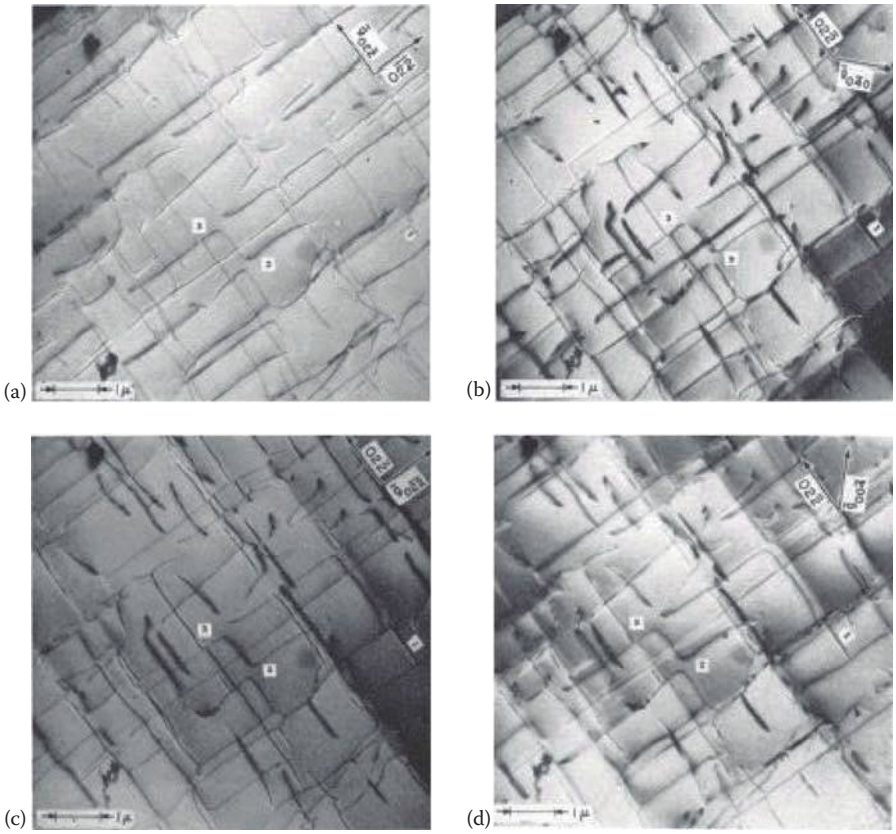


FIGURE 7.58
TEM Burgers vector analysis for a graded $GaAs_{1-x}P_x$ layer on a GaAs(001) substrate. The four plan-view TEM micrographs were obtained for the identical area of a single sample, but with different diffraction vectors: (a) $\mathbf{g} = [02\bar{2}]$, (b) $\mathbf{g} = [0\bar{2}\bar{2}]$, (c) $\mathbf{g} = [0\bar{4}0]$, and (d) $\mathbf{g} = [00\bar{4}]$. (Reprinted from M. S. Abrahams et al., *J. Mater. Sci.*, 4, 223 [1969]. With permission. Copyright 1969, Springer.)

reasoning, we can conclude that the misfit dislocation to the left of the numeral 1 is an edge dislocation, but with $\mathbf{b} = (a/2)[01\bar{1}]$.

In plan-view TEM micrographs like those shown in Figure 7.58, the points where misfit dislocations appear to end really correspond to their bending out of the plane of the image. This implies the existence of a set of threading dislocations, each of which has a component of its line vector perpendicular to the image plane. The Burgers vector is conserved upon bending of the dislocation, so an indirect Burgers analysis can be performed for the threading dislocations in this way. However, cross-sectional TEM micrographs are more commonly used for the study of threading dislocations, allowing their direct imaging and Burgers analysis.

7.11.2 Characterizing Defects by Crystallographic Etching

Crystallographic etches are often used to evaluate dislocation densities in bulk or heteroepitaxial semiconductors. These etches reveal the points of emergence of dislocations on the etched surface as hillocks or, more commonly, pits. These features occur due to the reduced or enhanced etch rate in the strained region around the dislocation.

Etch pits on a specimen surface are counted within a known area using OM, Nomarski phase contrast microscopy, SEM, AFM, or some other imaging method. The areal density of etch pits (the etch pit density [EPD]) can thus be determined, and is usually considered to correspond to the threading dislocation density.

The one-to-one correspondence between etch pits and dislocations was first established by Vogel et al.[135] They etched a bulk Ge crystal with 5:3:3:1/10 HNO_3:HF:CH_3COOH:Br_2 and showed that the spacing of the etch pits along a grain boundary was consistent with the angle of the boundary, as measured by XRD. However, each etchant and crystal system has unique characteristics, and it is always necessary to verify a one-to-one correspondence between etch pits and the threading dislocations using another technique, such as TEM.

In some cases, crystallographic etching results in distinctive etch pits that allow identification of different types of dislocations. An example of this is the case of offcut 4H- or 6H-SiC(0001) etched by molten KOH, in which three types of etch pits allow the identification of threading screw dislocations, threading edge dislocations, and basal-plane screw dislocations.[136,137]

Sometimes, the measured EPD is found to be much lower than the dislocation density determined by other techniques. In the case of GaAs/Si(001) etched by molten KOH, the measured etch pit densities are sometimes found to be orders or magnitude lower than the dislocation density measured by TEM.[138] This discrepancy has been explained by the existence of a second set of smaller etch pits that may be observed by TEM.[139] However, even the counting of both sets of pits resulted in underestimation of the dislocation density. This is because two or more closely spaced dislocations may produce a single pit, and because there are no pits associated with some dislocations.[140] Occasionally, a crystallographic etch will produce pits that do not appear to be associated with dislocations.

EPD measurements are best suited for crystals with low dislocation densities, such a semiconductor substrates. Because a large area can be etched and examined for pits, even Si wafers with dislocation densities of ~10 cm^{-2} may be characterized in this way. On the other hand, the etch pits from neighboring dislocations begin to overlap at high defect densities, resulting in undercounting. For this reason, the practical upper limit of the dislocation density that may be measured by crystallographic etching is about 10^6 to 10^8 cm^{-2}. This upper limit can be maximized by the use of slow etchants, short etch times, and high-magnification imaging techniques (SEM or TEM).

A number of crystallographic etches have been used with Si. The Dash etch[141] was first used to establish the correlation between etch pits and dislocations in Si; however, this mixture is slow and requires hours of etch time. The Sirtl,[142] Secco,[143] and Wright[144] etches are fast acting and commonly used with Si. All three of these etches utilize chromium salt oxidizers. Plating out of the chromium ion helps to delineate the etch pits under microscopic inspection. The Wright etch uses copper nitrate as a plating agent as well. The Sirtl and Wright etches are anisotropic, revealing triangular pits on Si(111) and rectangular pits on Si(001). The Secco etch is nearly isotropic, and produces circular or elliptical pits on both Si(111) and Si(001).

A wide range of crystallographic etches are available for GaAs as well. Molten KOH is commonly used, and an important application for this etch is the discovery of IDBs. The two-part A-B etch is frequently used for the determination of dislocation densities, but also for the delineation of interfaces in cleaved multilayered structures. In addition, all of the crystallographic etches designed for use with Si can also be used for GaAs. The application of Sirtl etch with GaAs is interesting in that it produces hillocks and mounds, rather than pits, at the emergence points for defects.

Molten KOH etching of GaAs was developed by Grabmaier and Watson[145] for use with the (001) orientation. Angilello et al.[146] subsequently showed the correspondence between etch pits revealed by molten KOH, and dislocations imaged in an x-ray transmission topograph using the 004 reflection, for GaAs(001). Typically, molten KOH etching is carried out in a crucible maintained at a temperature of 400°C or higher,* for a time of up to several minutes. On the (001) surface, the etch pits revealed by molten KOH elongate in the $[\bar{1}10]$ direction. They can therefore be used to distinguish between the $\langle 110 \rangle$ directions on the surface of a GaAs crystal. Moreover, in the presence of IDBs, the pits are rotated by 90° upon crossing a domain boundary. The presence (or absence) of these rotations is often used to establish whether inversion domains exist in the material.

Unfortunately, the crystallographic etches that were developed for GaAs seldom prove useful with other zinc blende III-V semiconductors, and so new etches have been developed and characterized for nearly every new material. In the case of InP, mixtures of HBr with either HF or CH_3COOH have been successfully applied to the EPD characterization of both (001) and (111) surfaces.[147] For GaP, H_2SO_4:H_2O_2:HF etches[148] are used. For InSb, HNO_3:HF:CH_3COOH etchant has been used for (001) and (111)Sb surfaces, but it is necessary to add stearic acid, Pb, or Ge in order to reveal etch pits on the (111)In surfaces.[149] Even for other arsenides, such as InGaAs, the A-B etch commonly employed for GaAs has proved to be unreliable in the measurement of dislocation densities.[150]

For zinc blende II-VI semiconductors as well, unique crystallographic etches have been developed for nearly every material of interest. Mixtures of lactic, nitric, and hydrofluoric acids are utilized with the tellurides CdTe[151] and CdZnTe,[152] whereas the Chen etch[153] is used with HgCdTe. Bromine–methanol mixtures have been used successfully with the selenides ZnSe,[122,154,155] ZnSSe, and ZnMgSSe.[156] This etch is ineffective with the cadmium-bearing quaternary ZnCdMgSe, however, so mixtures of hydrobromic and acetic acid are used instead.[157]

Several etches have been used for the EPD characterization of III-nitrides, including H_3PO_4:H_2SO_4 and molten KOH. Ono et al.[158] demonstrated the use of H_3PO_4:H_2SO_4 to reveal etch pits on GaN(0001), and this etch was used by Tsai et al.[159] to evaluate GaN free-standing films. Kozawa et al.[160] used molten KOH to characterize the dislocation density in GaN/α-Al_2O_3 (0001). They found hexagonal pits with a density of 2×10^7 cm^{-2}.

* The melting point for KOH is 360°C.

Rosner et al.[161] observed pits on the surface of GaN on sapphire (0001) following MOVPE growth, without the intentional use of a crystallographic etch. These pits were confirmed to be associated with threading dislocations. However, it was not clear whether the pits were inherent in the growth process or if they formed by an etching process in the reactor following growth.

SiC is difficult to etch due to its chemical stability, and molten KOH is the only reported crystallographic etch for this material. Typically, a temperature of 600°C is used with an etch time of 30 s to several minutes. Three types of etch pits are observed on vicinal 4H- or 6H-SiC(0001): large hexagonal pits, small hexagonal pits, and shell-like etch pits. These have been shown to be associated with threading screw dislocations, threading edge dislocations, and "basal-plane dislocations," respectively.[135,136] Threading screw dislocations and threading edge dislocations are parallel to the [0001]; the basal-plane dislocations are screw dislocations that lie within the (0001) basal plane but emerge at the surface if the SiC is offcut.

The compositions of some commonly used crystallographic etches are tabulated in Appendix VI.

7.11.3 Characterizing Defects by X-Ray Diffraction

The x-ray rocking curve technique is complementary to the TEM and EPD methods because it is nondestructive and can be used to determine threading dislocation densities in the range from 10^5 to 10^9 cm^{-2}, and with factor of two accuracy. Gay et al.[120] and Hordon and Averbach[121] described the theory of x-ray rocking curve broadening in dislocated metal crystals, and this work was extended to zinc blende semiconductor crystals.[121]

The x-ray rocking curve measured from a semiconductor crystal using a Bartels diffractometer or a double-crystal diffractometer in the parallel position has a characteristic FWHM that depends on the crystal being examined and the incident wavelength, but which is insensitive to instrumental effects. Dislocations broaden the rocking curve in three ways: (1) the dislocation introduces a rotation of the crystal lattice, thus directly broadening the rocking curve; (2) the dislocation is surrounded by a strained volume of crystal, in which the Bragg angle is nonuniform; and (3) in grossly dislocated crystals, arrays of dislocations can form the walls between small polycrystals, giving rise to crystal size broadening. In high-quality heteroepitaxial layers, only the first two (angular broadening and strain broadening) are important.

In developing a quantitative model for the dislocation broadening, it is assumed that the broadening is due to dislocation half loops or bent-over substrate threading dislocations. These dislocations have both threading and misfit dislocation segments that can give rise to rocking curve broadening. Kaganer et al.[173] have presented an analysis of the profile broadening due to misfit dislocations for certain defect configurations. However, the local strain variations introduced by misfit dislocations are expected to cancel at distances greater than one-half their mean spacing in the interface. For a relaxed layer, the misfit dislocation spacing in the interface is

$$S = \frac{1}{\rho} = \frac{b \sin\alpha \sin\phi}{|f|}, \tag{7.147}$$

where:

- f is the lattice mismatch strain
- b is the length of the Burgers vector
- α is the angle between the Burgers vector and line vector
- ϕ is the angle between the interface and the normal to the slip plane

For example, with $b = 4$ Å and $f = 0.1\%$, the misfit dislocation spacing is about 0.2 μm. A relaxed layer will be at least 10 times this thickness. Most of the diffracted intensity will originate in the top 90% of the layer, and the broadening will therefore be dominated by the threading dislocations there. Moreover, since the CLT and the misfit dislocation spacing in the relaxed layer are both inversely proportional to $|f|$, the same should be true for all relaxed heteroepitaxial layers.

Experimental results also suggest that the rocking curve broadening is dominated by threading dislocations. For example, Ayers et al.[162] showed that the misfit dislocations are not dominant in broadening the 004 rocking curves for GaAs/Si(001). In that study, GaAs-on-Si samples were prepared with equal misfit dislocation densities, but very different threading dislocation densities, by postgrowth annealing. It was found that the rocking curve widths changed dramatically with annealing, despite the essentially unchanged misfit dislocation densities.

It is further assumed that the resulting diffraction profiles are Gaussian. This assumption is supported by experimental evidence[160] in the case of GaAs/Si(001). Figure 7.59 shows the GaAs 004 rocking curve for a 1.5 μm thick GaAs/Si(001) sample grown by MOVPE, along with the Gaussian best fit. It can be seen that the two profiles are indistinguishable, except in the tail regions. This was also found to be true for the 002, 113, 224, 115, 006, 026, 444, and 117 rocking curves measured for the same GaAs/Si(001) sample. Kaganer et al.[163] found a similar result when they investigated the shapes of 0002 and 0004 diffraction profiles from dislocated GaN/6H-SiC(0001) grown by PAMBE. They found that the tails of the profiles obeyed a power law, but that most of the intensity resided in the central part of the peak, which was Gaussian in nature.

Assuming that the measured *hkl* rocking curve is Gaussian, with an FWHM $\beta_m(hkl)$, and results from the convolution of Gaussian intensity distributions,

$$\beta_m^2(hkl) = \beta_0^2(hkl) + \beta_d^2(hkl) + \beta_\alpha^2(hkl) + \beta_\varepsilon^2(hkl) + \beta_L^2(hkl) + \beta_r^2(hkl), \tag{7.148}$$

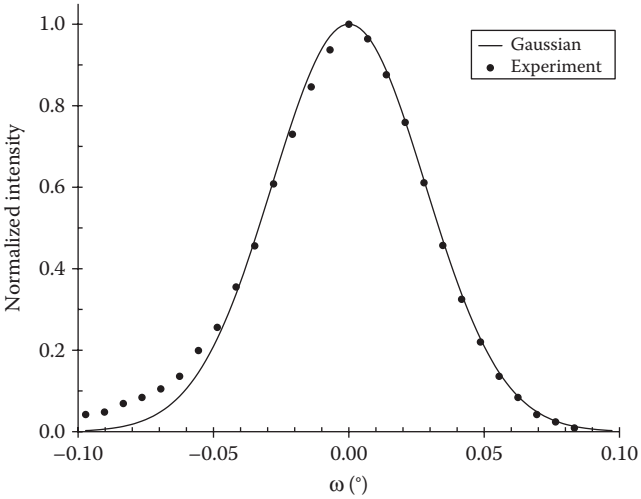


FIGURE 7.59
GaAs 004 rocking curve for a 1.5 μm thick GaAs/Si(001) sample grown by MOVPE, obtained by a double-crystal diffractometer with Cu kα radiation, along with the Gaussian best fit. (Adapted from J. E. Ayers, *J. Cryst. Growth*, 135, 71 [1994]. With permission. Copyright 1994, Elsevier.)

where:

- $\beta_0(hkl)$ is the intrinsic rocking curve width for the crystal being examined
- $\beta_d(hkl)$ is the width of the instrumental broadening function
- $\beta_\alpha(hkl)$ is the rocking curve broadening caused by angular rotations at dislocations
- $\beta_e(hkl)$ is the rocking curve broadening caused by the inhomogeneous strain surrounding dislocations
- $\beta_L(hkl)$ is the rocking curve broadening due to the crystal size (layer thickness)
- $\beta_r(hkl)$ is the rocking curve broadening due to curvature of the heteroepitaxial specimen

The intrinsic rocking curve width β_0 for $\theta_B + \phi$ incidence can be estimated by

$$\beta_0 = \frac{2r_e\lambda^2\left[1+\cos\left(2\theta_B\right)\right]|F_H|}{\pi V \sin\left(2\theta_B\right)}\left[\frac{\sin\left(\theta_B-\phi\right)}{\sin\left(\theta_B+\phi\right)}\right]^{1/2}, \tag{7.149}$$

where:

- r_e is the classical electron radius, 2.818×10^{-5} Å
- λ is the x-ray wavelength
- V is the crystal volume for which we have calculated the structure factor, $V = a_0^3$ for a cubic crystal
- $|F_H|$ is the magnitude of the structure factor for the *hkl* reflection
- θ_B is the Bragg angle

Selected values of these are tabulated in Appendix VII.

The broadening due to angular rotation at the dislocations has been modeled by Gay et al.[120] In this treatment, the single crystal is considered to comprise an arrangement of subsidiary mosaics with mutual inclination. Each subsidiary mosaic is associated with a dislocation, which tilts the mosaic with respect to its neighbors. If the orientations of the mosaics have a Gaussian distribution, then the probability that a mosaic is tilted by an angle θ from the mean orientation is equal to $(\sigma\sqrt{2\pi})^{-1}\exp(-\theta^2/2\sigma^2)$, where σ is the standard deviation. The probability for two mosaics to be tilted with respect to each other by an angle $|\theta-\phi|$ is equal to $(\sigma\sqrt{2\pi})^{-1}\exp[-(\theta^2+\phi^2)/2\sigma^2]$. The mean disorientation between mosaics is

$$\overline{|\theta-\phi|} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\theta-\phi|\exp\left[-\left(\theta^2+\phi^2\right)/\left(2\sigma^2\right)\right]d\theta\, d\phi}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\left(\theta^2+\phi^2\right)/\left(2\sigma^2\right)\right]d\theta\, d\phi}. \tag{7.150}$$

The value of the denominator is $2\pi\sigma^2$, and Dunn and Koch[164] have integrated the numerator explicitly to obtain $4\sigma^3\sqrt{\pi}$. Therefore,

$$\overline{|\theta-\phi|} = \frac{2\sigma}{\sqrt{\pi}}. \tag{7.151}$$

For a Gaussian distribution, the relationship between the standard deviation σ and the FWHM β is given by $\sigma = \beta/(2\sqrt{2\ln 2})$, so that the mean disorientation between mosaics can be related to the measured FWHM by

$$\overline{|\theta-\phi|}=\frac{\beta_\alpha}{\sqrt{2\pi\ln 2}}. \tag{7.152}$$

If the threading dislocations are arranged in a random network, with an average spacing of $1/\sqrt{D}$, where D is the dislocation density, then the average angle between neighboring mosaics is approximately

$$\overline{|\theta-\phi|}\approx b\sqrt{D}, \tag{7.153}$$

where b is the length of the Burgers vector. With Equations 7.152 and 7.153, we can relate the measured broadening to the dislocation density by

$$D\approx\frac{\beta_\alpha^2}{(2\pi\ln 2)b^2}=\frac{\beta_\alpha^2}{4.36b^2}. \tag{7.154}$$

The angular broadening due to threading dislocations is therefore given by

$$\beta_\alpha^2\approx(2\pi\ln 2)b^2D=K_\alpha. \tag{7.155}$$

The strain broadening due to dislocations has been modeled by Warren[165] and Hordon and Averbach.[120] If it is assumed that the random array of threading dislocations gives rise to a Gaussian distribution of local strain, then this strain gives rise to rocking curve broadening given by

$$\beta_\varepsilon^2=(8\ln 2)\overline{\varepsilon_N^2}\tan^2\theta_B=K_\varepsilon\tan^2\theta_B, \tag{7.156}$$

where $\overline{\varepsilon_N^2}$ is the mean square strain in the direction of the normal to the diffracting planes. $\overline{\varepsilon_N^2}$ may be estimated as follows. If it is assumed that the Poisson ratio is 1/3, then from the known strain distribution around the edge dislocation it is found that

$$\overline{\varepsilon_{Ne}^2}=\frac{5b^2\left(2.45\cos^2\Delta+0.45\cos^2\psi\right)}{64\pi^2r^2}\ln\left(\frac{r}{r_0}\right), \tag{7.157}$$

where:

- Δ is the angle between the dislocation glide plane normal and the normal to the diffracting planes
- ψ is the angle between the dislocation Burgers vector and the normal to the diffracting planes
- r and r_0 are the upper and lower limits for integration of the strain field in the radial direction from the dislocation core

Similarly, for the pure screw dislocation,

$$\overline{\varepsilon_{Ns}^2}=\frac{b^2\sin^2\psi}{4\pi^3r^2}\ln\left(\frac{r}{r_0}\right). \tag{7.158}$$

For dislocations with mixed character,

$$\overline{\varepsilon_N^2} = \overline{\varepsilon_{Ne}^2}\sin^2\alpha + \overline{\varepsilon_{Ns}^2}\cos^2\alpha\,, \tag{7.159}$$

where α is the angle between the Burgers vector and line vector for the dislocations.

For half loops in a (001) diamond or zinc blende epitaxial layer, with 60° misfit segments and screw (90°) threading segments, it has been shown that the dislocation broadening is given by

$$\begin{aligned}\beta_\varepsilon^2 &= 0.09 b^2 D \left|\ln\left(2\times10^{-7}\, cm\sqrt{D}\right)\right| \tan^2\theta_B \\ &= K_\varepsilon \tan^2\theta_B.\end{aligned} \tag{7.160}$$

For heteroepitaxial layers, the crystal size broadening is usually affected only by the layer thickness *h*. Then, the crystal size broadening is given approximately by the Scherrer equation:[113]

$$\beta_L^2 \approx \left[\frac{4\ln 2}{\pi h^2}\right]\left(\frac{\lambda^2}{\cos^2\theta_B}\right). \tag{7.161}$$

For example, for the 004 reflection from GaAs(001), $\beta_L = 35''$ for a 1 μm thick layer.

Rocking curve broadening due to specimen curvature has been analyzed by Halliwell et al.[25] and Flanagan.[166] It is given by

$$\beta_r^2 = \frac{w^2}{r^2\sin^2\theta_B}, \tag{7.162}$$

where:

w is the width of the x-ray beam in the plane of the diffractometer
r is the radius of curvature for the heteroepitaxial structure

Usually, the broadening due to curvature is negligible if the substrate is much thicker than the epitaxial layer. However, since the curvature contribution is the same for both the substrate and the epitaxial layer, the substrate rocking curve width gives an upper limit for the broadening due to the curvature.

Combining the above equations, we obtain

$$\begin{aligned}\beta_m^2(hkl) &= \beta_0^2(hkl) + \beta_d^2(hkl) + K_\alpha + K_\varepsilon \tan^2\theta_B \\ &\quad + \left[\frac{4\ln 2}{\pi h^2}\right]\left(\frac{\lambda^2}{\cos^2\theta_B}\right) + \frac{K_r}{\sin^2\theta_B}.\end{aligned} \tag{7.163}$$

If the effects of the crystal size broadening and curvature are negligible, then the broadening contribution due to dislocations β_{disl} can be found by the deconvolution of the natural width of the reflection and the instrumental broadening:

$$\beta_{\text{disl}}^2 = \beta_m^2(hkl) - \beta_d^2(hkl) = K_\alpha + K_\varepsilon \tan^2\theta_B. \tag{7.164}$$

If the rocking curve width $\beta_m(hkl)$ is measured for a number of *hkl* reflections, and the extracted value $\beta_{\text{disl}}(hkl)$ is plotted as a function of $\tan^2\theta_B$, then the values of K_α and K_ε correspond to the intercept and slope of the characteristic, respectively. The dislocation density can be found from either value, using the angular broadening,

$$D = \frac{K_\alpha}{4.36b^2}, \tag{7.165}$$

or the strain broadening,

$$D = \frac{K_\varepsilon}{0.09b^2\left|\ln\left(2\times10^{-7}\,\text{cm}\sqrt{D}\right)\right|}. \tag{7.166}$$

Figure 7.60 shows the application of this method, for the case of a 3.0 μm thick layer of GaAs/Si(001) grown by MOVPE. The filled circles represent measurements and the line is the least-squares fit, given by

$$\beta_{\text{disl}}^2 = 41{,}600\ (\text{arc sec})^2 + 5{,}850\ (\text{arc sec})^2 \tan^2\theta_B. \tag{7.167}$$

Therefore, from the angular broadening of dislocations we obtain $D = 1.4 \times 10^8$ cm^{-2} and from the strain broadening of the dislocations we obtain $D = 1.5 \times 10^8$ cm^{-2}.

Because the two approaches give nearly the same dislocation density, it is usually adequate to estimate D from the angular broadening alone. Then it suffices to measure one rocking curve at a small value of $\tan^2\theta_B$. In the case of GaAs(001), the 004 reflection can be used because for this reflection, $\tan^2\theta_B = 0.422$, so the strain broadening may be neglected.

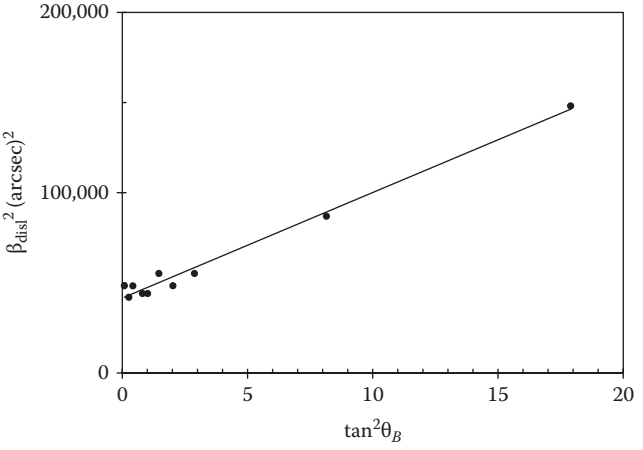


FIGURE 7.60
β_{disl}^2 vs. $\tan^2\theta_B$ for a 3.0 mm thick layer of GaAs/Si(001) grown by MOVPE. β_{disl}^2 is the square of the dislocation broadening, extracted from measured rocking curve widths for various *hkl* reflections. θ_B is the Bragg angle. The filled circles represent the data extracted from measurements, and the line is the least-squares fit, given by $\beta_{\text{disl}}^2 = 41{,}600(\text{arc sec})^2 + 5{,}850(\text{arc sec})^2 \tan^2\theta_B$. (Adapted from J. E. Ayers, *J. Cryst. Growth*, 135, 71 [1994]. With permission. Copyright 1994, Elsevier.)

7.11.4 Characterization of Asymmetric Dislocation Densities

For (001) heteroepitaxy of zinc blend semiconductors, the misfit dislocation densities along [110] and $[1\bar{1}0]$ directions may be different. Such an asymmetry may arise as a consequence of the difference in mobilities for α dislocations (aligned with the $[1\bar{1}0]$ direction) and β dislocations (aligned with the [110] direction) in the semiconductor.[167] This imbalance in dislocation populations is reflected in the densities of threading dislocations on the two types of glide systems and may be characterized by observing the variation of the x-ray rocking curve width with azimuthal direction. In order to apply this technique, it is necessary to distinguish between the [110] and $[1\bar{1}0]$ directions of the crystal, but this can be done by applying a crystallographic etch to the surface. Typically, such an etch reveals pits that are elongated in the [110] direction. Yarlagadda et al.[168] have described this characterization method along with a quantitative model for its analysis.

For the heteroepitaxial growth of a zinc blende semiconductor with the (001) orientation, there are eight active slip systems, as given in Table 7.2.[167] We can further group these slip systems in two groups based on the orientation of the line vectors for their misfit dislocations. Slip systems with their misfit dislocation line vectors along the $[1\bar{1}0]$ are considered of type A, associated with α-type misfit segments, whereas those with their misfit dislocation line vectors along the [110] are considered type B, associated with β-type misfit segments. The densities of these dislocations are related by

$$D_A = D_1 + D_2 + D_3 + D_4 \tag{7.168}$$

and

$$D_B = D_5 + D_6 + D_7 + D_8. \tag{7.169}$$

TABLE 7.2

Active Slip Systems for (001) Heteroepitaxy of Diamond or Zinc Blende Semiconductors

Slip System	Misfit Dislocation Line Vector	Glide Plane	Burgers Vector b	TD Line Vector l
S_1	$[1\bar{1}0]$	(111)	$\frac{1}{2}a[10\bar{1}]$	$\frac{1}{2}a[0\bar{1}1]$
S_2	$[1\bar{1}0]$	(111)	$\frac{1}{2}a[01\bar{1}]$	$\frac{1}{2}a[\bar{1}01]$
S_3	$[1\bar{1}0]$	$(\bar{1}\bar{1}1)$	$\frac{1}{2}a[101]$	$\frac{1}{2}a[011]$
S_4	$[1\bar{1}0]$	$(\bar{1}\bar{1}1)$	$\frac{1}{2}a[011]$	$\frac{1}{2}a[101]$
S_5	[110]	$(\bar{1}11)$	$\frac{1}{2}a[101]$	$\frac{1}{2}a[0\bar{1}1]$
S_6	[110]	$(\bar{1}11)$	$\frac{1}{2}a[01\bar{1}]$	$\frac{1}{2}a[101]$
S_7	[110]	$(1\bar{1}1)$	$\frac{1}{2}a[10\bar{1}]$	$\frac{1}{2}a[011]$
S_8	[110]	$(1\bar{1}1)$	$\frac{1}{2}a[011]$	$\frac{1}{2}a[\bar{1}01]$

If symmetric rocking curves are measured at a Bragg angle θ_B such that $\tan^2\theta_B < 5$ (such as 004 rocking curves measured with Cu k_{a1} radiation) and the dislocation density is greater than about 10^6 cm^{-2}, then the rocking curve width is dominated by the angular mosaic spread introduced by dislocations.[121] The type A and type B dislocations will broaden the symmetric rocking curve when the azimuth is such that the projection of the incident beam is parallel to the misfit dislocation line vector, but not when these directions are normal.

If the epitaxial layer is laterally uniform and grown upon an on-axis substrate, and the incident x-ray spot is aligned with the axis of azimuthal rotation, then a variation of the rocking curve width with azimuth indicates an imbalance in the dislocation populations on the type A and type B slip systems. The individual dislocation populations can be estimated from the measured rocking curve width dependence on the azimuth ψ. Suppose the zero azimuth ($\psi = 0$) corresponds to the case with the projection of the incident beam aligned with the [110] direction in the surface of the sample and ($\psi = \underline{9}0°$) corresponds to the case with the projection of the incident beam aligned with the [1$\bar{1}$0] direction in the surface of the sample. Suppose γ_A and γ_B are the angles between the Burgers vectors and the diffraction vectors for dislocations on the type A and type B slip systems, and b is the length of the Burgers vector. Assuming the applicability of the mosaic block model, the mean misorientation of crystal blocks about an axis in the plane of the interface and perpendicular to the incident x-ray beam, associated with dislocations on type A slip systems with a density D_A, will be $2b\sqrt{D_A}\sin(\psi-\psi_A)\cos\gamma_A$. Assuming a Gaussian distribution of the crystal block orientations, the x-ray rocking curve width β_A associated with dislocations on type A slip systems is given by

$$\beta_A^2 = 8\pi D_A \ln(2) b^2 \sin^2(\psi-\psi_A)\cos^2\gamma_A, \tag{7.170}$$

where $\psi_A = 90°$. Similarly, for dislocations on type B slip systems,

$$\beta_B^2 = 8\pi D_B \ln(2) b^2 \sin^2(\psi-\psi_B)\cos^2\gamma_B, \tag{7.171}$$

where $\psi_B = 0°$. To find the overall rocking curve width, it is assumed that the broadening functions convolve approximately like Gaussian distributions—a reasonable assumption as long as the rocking curve is verified to have an approximately Gaussian shape. Then, the square of the rocking curve width is given by

$$\beta^2 = \beta_A^2 + \beta_B^2, \tag{7.172}$$

and the rocking curve width β will vary with azimuth as

$$\beta^2 = 8\pi\ln(2) b^2\left[\cos^2\psi\cos^2\gamma_A D_A + \sin^2\psi\cos^2\gamma_B D_B\right]. \tag{7.173}$$

The individual dislocation densities on the two types of slip systems can be estimated from

$$D_A \approx \frac{\beta^2(\psi=0)}{8\pi\ln(2) b^2\cos^2\gamma_A} = \frac{\beta^2(\psi=0)}{8.7b^2} \tag{7.174}$$

and

$$D_B \approx \frac{\beta^2\left(\psi = 90^\circ\right)}{8\pi \ln(2) b^2 \cos^2 \gamma_B} = \frac{\beta^2\left(\psi = 90^\circ\right)}{8.7 b^2}. \tag{7.175}$$

Yarlagadda et al.[168] applied this technique to ZnSe/GaAs(001) and ZnS_ySe_{1-y}/GaAs(001) heterostructures grown by photoassisted MOVPE.[169,170] In that work, the actual misorientation of the substrates was reported to be less than 0.25° from the XRD results. The [110] and [1$\bar{1}$0] directions were distinguished by employing a 1:400 bromine–methanol etch for 16 s at 0°C. This crystallographic etch reveals pyramidal pits that elongate in the [110] direction,[154] and which were observed by OM at 1000× magnification, as shown in Figure 7.61. Symmetric 004 rocking curves were measured at a number of different azimuths using Cu kα_1 radiation and a Bartels-type diffractometer. Figure 7.62 shows the 004 rocking curve

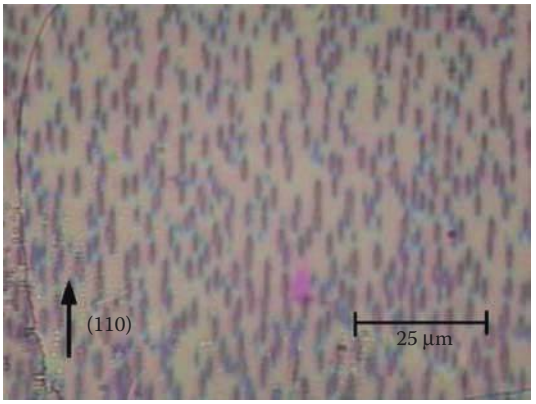


FIGURE 7.61
Etch pits observed on a ZnSe/GaAs(001) epitaxial layer after etching for 6 s at 0°C in 1:400 bromine–methanol. (Reprinted from B. Yarlagadda et al., *Appl. Phys. Lett.*, 92, 202103 [2008]. With permission. Copyright 2008, American Institute of Physics.)

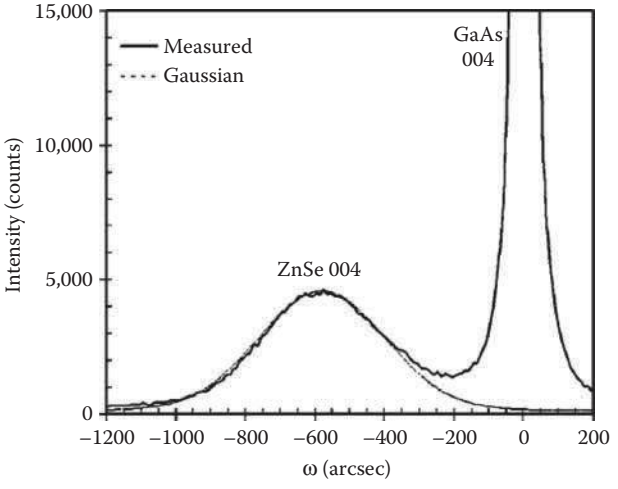


FIGURE 7.62
004 X-ray rocking curve for ZnSe/GaAs(001) epitaxial layer. (Reprinted from B. Yarlagadda et al., *Appl. Phys. Lett.*, 92, 202103 [2008]. With permission. Copyright 2008, American Institute of Physics.)

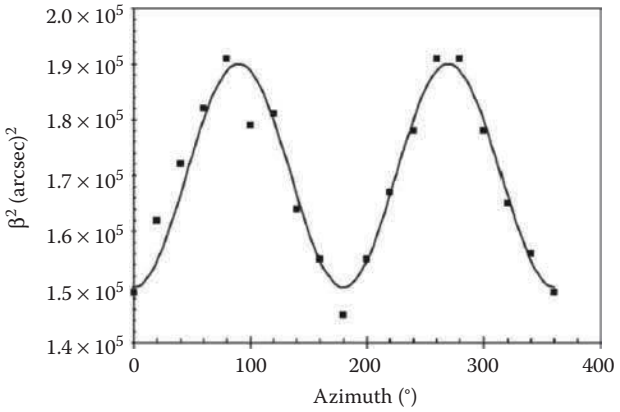


FIGURE 7.63
Square of the 004 rocking curve FWHM as a function of the azimuthal angle for a ZnSe/GaAs(001) epitaxial layer. (Reprinted from B. Yarlagadda et al., *Appl. Phys. Lett.*, 92, 202103 [2008]. With permission. Copyright 2008, American Institute of Physics.)

for this sample at zero azimuth, along with a Gaussian best fit for the ZnSe 004 diffraction peak. The square of the rocking curve FWHM was plotted as a function of the azimuthal angle and is shown in Figure 7.63. The filled squares show experimental data, and the curve shows the fit to Equation 7.173 given by

$$\left(\beta/\text{arc sec}\right)^2 = 1.5\times10^5\cos^2\psi + 1.9\times10^5\sin^2\psi. \tag{7.176}$$

This variation was attributed to the asymmetry in dislocation densities on the two types of slip systems after ruling out other possible causes, and the two densities were estimated to be $D_A \approx 2.5 \times 10^8$ cm^{-2} and $D_B \approx 2.0 \times 10^8$ cm^{-2}. Although this quantitative analysis is approximate, the results have good relative accuracy, so it was concluded that D_A is ~25% higher than D_B.

Yarlagadda et al.[167] also observed asymmetric dislocation densities in a 2 μm $ZnS_{0.02}Se_{0.98}$ device layer grown on a retro-graded ZnS_ySe_{1-y} buffer. The retro-graded buffer was 0.7 μm thick and graded from y = 0 at the bottom to y = 0.04 at the top. (This results in retro grading because the lattice matching condition for ZnS_ySe_{1-y}/GaAs(001) occurs with a sulfur mole fraction of ~6%.) The 004 rocking curve, with most intensity originating from the top device layer, exhibited a variation in the 004 rocking curve width given by

$$\left(\beta/\text{arc sec}\right)^2 = 1.4\times10^5\cos^2\psi + 2.2\times10^5\sin^2\psi. \tag{7.177}$$

For a graded structure such as this one, an additional complication arises because there might be nonuniform crystallographic tilting of the graded layer with distance from the interface.[106] However, this could be ruled out because such a nonuniform tilting would cause a variation in the symmetry of the rocking curve, as well as its width, but this was not found. Therefore, the dislocation densities for the top device layer could be estimated as $D_A \approx 3.7 \times 10^8$ cm^{-2} and $D_B \approx 2.3 \times 10^8$ cm^{-2}, so for the retro-graded sample, D_A is ~60% higher than D_B.

7.11.5 Diffuse Scattering Characterization of Defects

Koppensteiner et al. have shown that the threading dislocation density may be determined from the diffuse scattering by using the two-dimensional intensity distribution from an RSM.[171] In this indirect approach, the dislocations are assumed to introduce random elastic deformations, and the resulting RSM contours are simulated to determine the energy density stored in the heteroepitaxial layer by the dislocations. The threading dislocation density is obtained by dividing this energy density by the line energy for 60° dislocations. This approach, while similar in principle to the approach based on the x-ray FWHM, is more difficult to apply because the details of the RSM simulations were not given.

7.11.5.1 Application of X-Ray Reciprocal Space Maps to Characterize Defects in IMF-Grown Materials

An interesting application of reciprocal space mapping is the characterization of high-mismatch materials grown in an interfacial misfit (IMF) dislocation growth mode,[172] whereby all of the mismatch strain is (hopefully) relaxed by the introduction of an ideal array of edge misfit dislocations, in the absence of threading dislocations. Here, the proposed method involves a comparison of the ratio of the x-ray FWHMs from the RSM, $\mathrm{FWHM}_x/\mathrm{FWHM}_z$, with the theoretical value, where FWHM_x is the width for the in-plane direction and FWHM_z is the width for the out-of-plane (growth) direction. Kaganer et al.[173] have calculated the theoretical RSM profile for materials grown in an ideal IMF mode with a perfect array of edge dislocations at the interface and absent of all threading dislocations. The RSM is expected to exhibit an elliptical peak with intensity proportional to

$$I(q_x, q_z) \infty \exp\left[-\frac{x_0 (q_x - q_{x_0})^2}{4\sigma_x^2(0)} - \frac{(q_z - q_{z_0})^2}{4\sigma_z^2(0)} \right], \tag{7.178}$$

where $(q_x - q_{x_0})$ and $(q_z - q_{z_0})$ are the deviations from the reciprocal space point (q_{x_0}, q_{z_0}) in the in-plane and out-of-plane directions, respectively, and the standard deviations are given by

$$\sigma_x^2(0) = \frac{\rho Q_z^2 b_x^2}{8\pi h} \tag{7.179}$$

and

$$\sigma_z^2(0) = \left(\frac{\nu}{1-\nu} \right)^2 \frac{\rho Q_z^2 b_x^2}{4\pi h}, \tag{7.180}$$

where:

- ρ is the linear density of misfit dislocations
- ν is the Poisson ratio
- Q_z is the component of the reciprocal space vector in the growth direction
- b_x is the in-plane component of the Burgers vector for the edge dislocations
- h is the film thickness

The ratio of the widths in this ideal IMF case is therefore predicted to be

$$\frac{\mathrm{FWHM}_x}{\mathrm{FWHM}_z} = \frac{\sigma_x(0)}{\sigma_z(0)} = \frac{1-\nu}{\sqrt{2}\nu}, \tag{7.181}$$

which is equal to 1.574 for GaSb/GaAs(001), 1.313 for InAs/GaAs(001), and 1.257 for InP/GaAs(001). Whereas the absolute widths depend on the film thickness, the ratio of widths does not. Figure 7.64 shows the 004 XRD profile, AFM topograph, and x-ray RSM

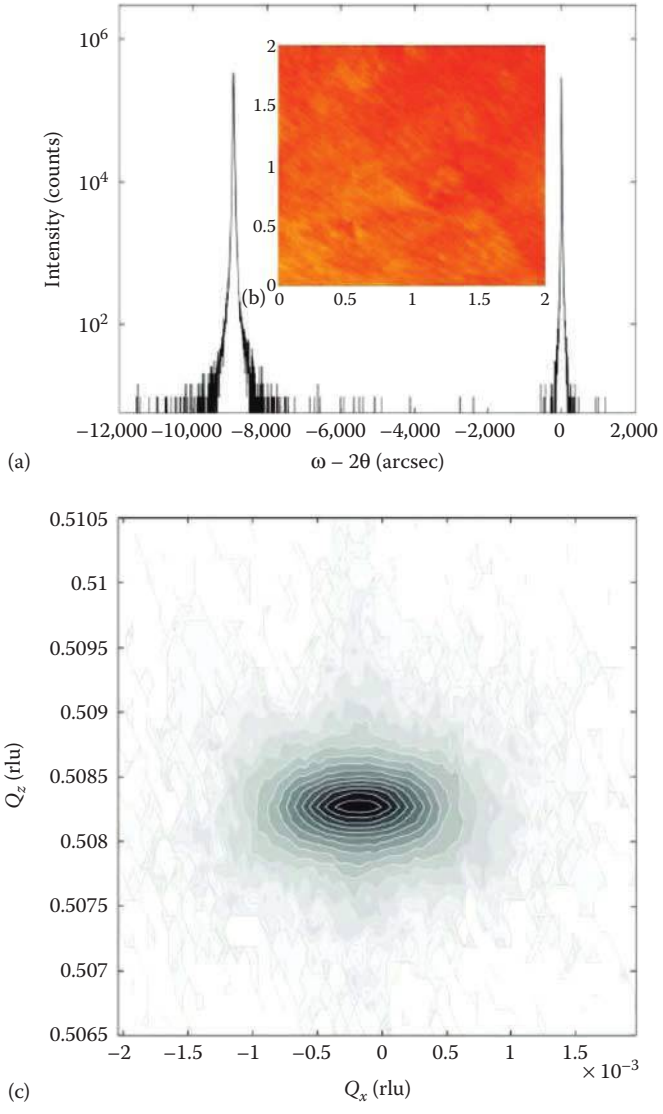


FIGURE 7.64
(a) 004 XRD profile for a 5 μm thick layer of InAs on GaAs(001), showing an FWHM of 28 arcsec. (b) AFM topograph for the same film, for which the horizontal scan direction is [110]. (c) RSM around the 004 peak of the GaSb, showing the asymmetry in the in-plane (Q_x) and out-of-plane (Q_z) directions. (Reprinted from C. J. K. Richardson et al., *J. Vac. Sci. Technol. B*, 29, 03C126-1 [2011]. With permission. Copyright 2011, American Vacuum Society.)

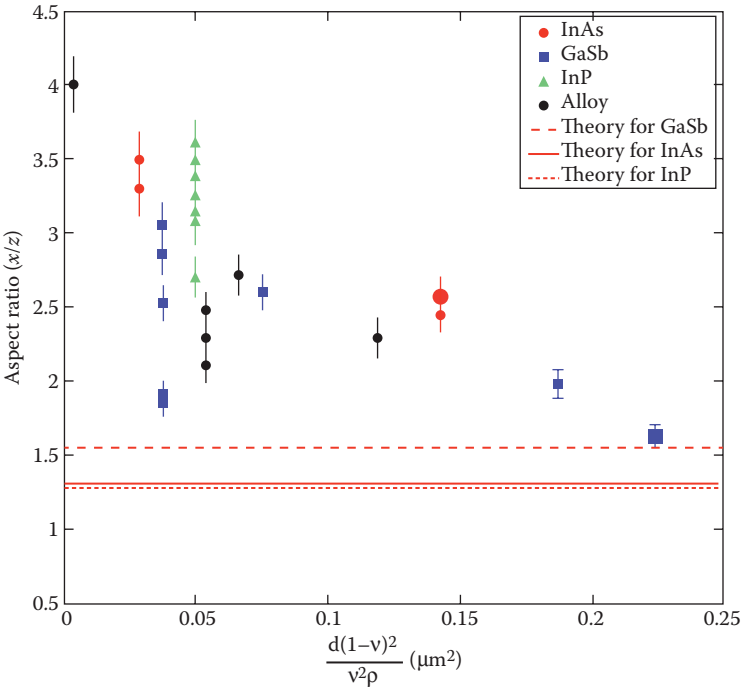


FIGURE 7.65
Comparison of experimental FWHM ratios from the RSM with the theoretical values for GaSb(1.574), InAs(1.313), and InP(1.257) on GaAs(001) substrates. The points labeled "alloy" were for ternary films containing various combinations of the atoms Al, Ga, In, As, and Sb. (Reprinted from C. J. K. Richardson et al., *J. Vac. Sci. Technol. B*, 29, 03C126-1 [2011]. With permission. Copyright 2011, American Vacuum Society.)

around the GaSb 004 reciprocal space point for a 5 μm thick InAs layer on GaAs(001). This sample exhibits a width ratio of $FWHM_x/FWHM_z = 2.6$, which is within a factor of two of the theoretical value.

Richardson et al. also plotted experimental ratios of $FWHM_x/FWHM_z$ for a number of metamorphic III-V materials on GaAs substrates and compared them with the theoretical values for GaSb, InAs, and InP, as shown in Figure 7.65. Although GaSb/GaAs(001), the prototypical material system for IMF growth, exhibits an FWHM ratio that is equal to the theoretical value, the other materials studied did not. This departure may be caused by the presence of threading dislocations, or it could be due to the presence of dislocations having mixed character in the interfacial array.

7.12 Dynamical X-Ray Diffraction Analysis of Multilayered Device Structures and Superlattices

For multilayered heteroepitaxial structures, a wealth of information can be obtained from the x-ray rocking curve, including the depth profiles of composition and strain. However, the analysis is not straightforward, as in the case of a single, uniform heteroepitaxial layer. Tanner and Hill[174] have shown that there is no one-to-one correspondence between the

layers in the structure and the intensity peaks in the rocking curve. Moreover, because the phase information is lost in the x-ray rocking curve, it is not possible to directly calculate the structure from it. Instead, this is done *indirectly* with dynamical x-ray simulations. The starting point for this analysis is a virtual structure, which represents an educated guess based on the growth conditions and times. The virtual structure is then refined by adjusting the thicknesses, compositions, or lattice constants in the individual layers until there is an acceptable match between the simulated rocking curve and the experimental profile. The refined virtual structure is then assumed to correspond closely to the actual one. This analysis requires the ability to calculate the x-ray rocking curve for an arbitrary layered structure, which will be described briefly here.

The calculation of the rocking curve for an arbitrary structure is based on recursion formulae derived from the Tagaki–Taupin equation,[25,26,175] which was introduced in Section 7.2.3 and is repeated here:

$$-i\frac{dX}{dT} = X^2 - 2\eta X + 1. \tag{7.182}$$

The structure is assumed to comprise N uniform layers, as shown in Figure 7.66. This approach is generally applicable, and continuously graded layers may be approximated by a series of steps.

The starting point is the calculation of the scattering amplitude for the substrate, which is assumed to be an infinitely thick, perfect crystal. It is calculated using the Darwin–Prins formula:[33]

$$X = \eta_S - \mathrm{Sign}(\eta_S)\sqrt{\eta_S^2 - 1}, \tag{7.183}$$

where the deviation parameter for the substrate is given by

$$\eta_S = \frac{-b(\theta - \theta_{BS})\sin(2\theta_{BS}) - 0.5(1-b)\Gamma F_{0S}}{\sqrt{|b|}C\Gamma\sqrt{F_{HS}F_{\bar{H}S}}}, \tag{7.184}$$

where:

b is the asymmetry factor, given by the ratio of the direction cosines for the incident and diffracted beams

θ_{BS} is the Bragg angle for the substrate

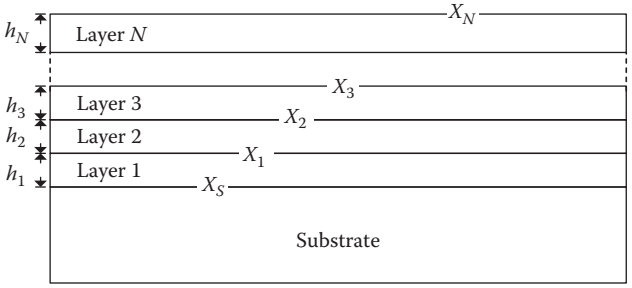


FIGURE 7.66
Lamellar (layered) structure assumed for the dynamical simulation of the x-ray rocking curve. The substrate is assumed to be an infinitely thick, perfect crystal. The N layers are assumed to be uniform both laterally and in the growth direction.

θ	is the actual angle of incidence on the diffracting planes
F_{0s}, F_{HS}, and $F_{\bar{H}\bar{S}}$	are the substrate structure factors for the 000, *hkl*, and $\bar{h}\bar{k}\bar{l}$ reflections, respectively
C	is the polarization factor
Γ	is given by

$$\Gamma = \frac{r_e \lambda^2}{\pi V}, \tag{7.185}$$

where:

r_e	is the classical electron radius, 2.818×10^{-5} Å
λ	is the x-ray wavelength
V	is the crystal volume for which we have calculated the structure factor (For a cubic crystal, $V = a_0{}^3$.)

When the polarization of the incident beam is in the plane of incidence (π polarization), $C_\pi = |\cos(2\theta_B)|$, and when the x-rays are polarized perpendicular to the plane of incidence (σ polarization), $C_\sigma = 1$.

Once the scattering amplitude has been calculated for the substrate, recursion equations are used N times for the N layers in the virtual structure. For the nth layer in the stack, the scattering amplitude at the top X_n is related to the scattering amplitude at the bottom X_{n-1} by

$$X_n = \eta_n + \sqrt{\eta_n^2 - 1}\,\frac{(S_{1n} + S_{2n})}{S_{1n} - S_{2n}}, \tag{7.186}$$

where

$$S_{1n} = \left(X_{n-1} - \eta_n + \sqrt{\eta_n^2 - 1}\right)\exp\left(-iT_n\sqrt{\eta_n^2 - 1}\right) \tag{7.187}$$

and

$$S_{2n} = \left(X_{n-1} - \eta_n - \sqrt{\eta_n^2 - 1}\right)\exp\left(iT_n\sqrt{\eta_n^2 - 1}\right). \tag{7.188}$$

Here, the deviation parameter is the value for the nth layer, given by

$$\eta_n = \frac{-b(\theta - \theta_{Bn})\sin(2\theta_{Bn}) - 0.5(1-b)\Gamma F_{0n}}{\sqrt{|b|}\,\Gamma\sqrt{F_{Hn}F_{\bar{H}n}}}, \tag{7.189}$$

and the thickness parameter T_n for the nth layer is

$$T_n = h_n\frac{\pi\Gamma\sqrt{F_{Hn}F_{\bar{H}n}}}{\lambda\sqrt{|\gamma_0\gamma_H|}}, \tag{7.190}$$

where h_n is the thickness of the nth layer and F_{0n}, F_{Hn}, and $F_{\bar{H}n}$ are the structure factors for the nth layer.

The substrate scattering amplitude is used as the scattering amplitude at the bottom of layer 1, for the calculation of the scattering amplitude at the top of layer 1. This is the scattering amplitude at the bottom of layer 2, which is used to obtain the scattering amplitude at the top of that layer, and so on. Once the scattering amplitude X_N has been calculated for the top layer of the stack, the reflectivity is calculated from

$$P_H = \left|\frac{F_{HN}}{F_{\overline{H}N}}\right| |X_N|^2. \tag{7.191}$$

This process is repeated for each angle in the range of interest to obtain the simulated rocking curve.

Figure 7.67 shows the measured and simulated x-ray rocking curves for a 10-period InGaAs/InP superlattice grown by MOVPE.[176] The various orders of the superlattice peaks are indexed in the figure. The InGaAs and InP layers in the superlattice have thicknesses of 2 and 140 nm, respectively. It can be seen that the simulated rocking curve matches the measured profile very well, except in the low-intensity troughs where the experimental result is dominated by noise. In order to obtain good agreement with the measured rocking curve, it was necessary to assume the presence of a $InAs_xP_{1-x}$ graded layer at each interface where InP was grown after InGaAs. The graded layer was assumed to have an arsenic composition x that decreased exponentially from 7.4% at the nominal interface, with a characteristic depth of 18 nm. In this case, therefore, the rocking curve analysis was used to deduce this departure from the ideal growth structure.

Whereas the rocking curve simulation approach is generally applicable to heteroepitaxial multilayers, superlattice structures represent a special case for which the superlattice period can be estimated without the need for rocking curve simulations. If the number of periods in the superlattice exceeds the number of atomic layers within each period, the

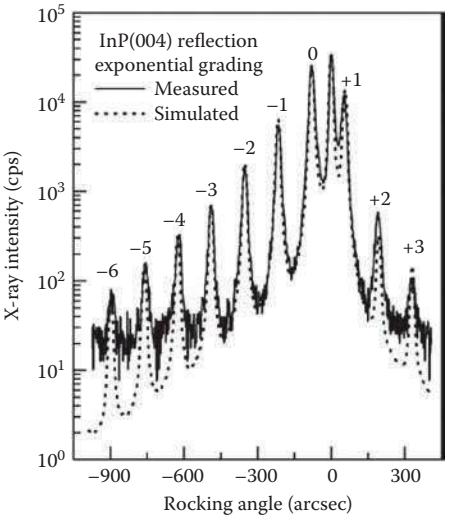


FIGURE 7.67
Measured and simulated x-ray rocking curves for a 10-period InGaAs/InP superlattice. The various orders of the superlattice peaks are indexed in the figure. The InGaAs and InP layers in the superlattice have thicknesses of 2 and 140 nm, respectively. (Reprinted from I. Kim et al., *J. Appl. Phys.*, 83, 3932 [1998]. With permission. Copyright 1998, American Institute of Physics.)

interference peaks in the tales of the rocking curve will have a period dominated by the superlattice. Then, if the superlattice period is Λ, the angular spacing of the superlattice peaks in the rocking curve will be

$$\Delta\theta = \frac{\lambda}{\Lambda\cos(\theta_B)}. \tag{7.192}$$

For example, using the 004 reflection for an AlAs/GaAs superlattice with $\lambda \approx 1.540594$Å and $\theta_B \approx 33.0°$, a superlattice period of 50 nm will result in superlattice peaks separated by ~760 arcsec.

Fewster et al.[177] have extended the dynamical simulation of XRD profiles to InGaN/GaN structures. The basic approach is similar to the case of cubic semiconductors, as long as the correct structure factors and geometric factors are applied. This is illustrated in Figure 7.68, which shows the simulated and experimental 0002 diffraction profiles for a single 2.7 nm $In_{0.13}Ga_{0.87}N$/GaN quantum well. For a quantum well structure such as this, good agreement is obtained between the dynamical calculation and the experiment without accounting for the effects of threading dislocations.

7.12.1 Dynamical X-Ray Diffraction Analysis of Device Heterostructures Containing Dislocations

HRXRD measurements allow determination of depth profiles of strain, composition, and defect densities in device structures, as described previously. Dynamical[22–26] and kinematical[9,12,14,118,178] simulations have been used in conjunction with a curve-fitting procedure to extract the profiles of strain and composition, but are based on perfect, dislocation-free laminar crystals, which renders the analysis inapplicable to mismatched structures

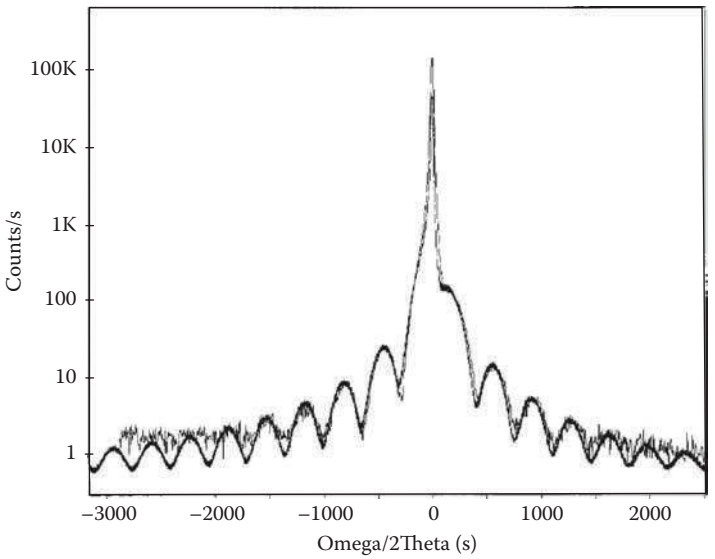


FIGURE 7.68
ω/2θ diffraction profiles for the 0002 reflection from a single 2.7 nm $In_{0.13}Ga_{0.87}N$/GaN quantum well. The heavy line corresponds to the dynamical calculation, while the light line corresponds to the experimentally measured profile. (Reprinted from P. F. Fewster et al., *J. Vac. Sci. Technol. B*, 18, 2300 [2000]. With permission. Copyright 2000, American Vacuum Society.)

with dislocation densities greater than 10^6 cm^{-2}. Dislocation densities much higher than this are encountered in metamorphic device structures, so that the effects of the dislocations on the rocking curves must be considered. Krivoglaz and Ryaboshapka[179] and Levine and Thomson[180] have analyzed the line profiles of Bragg peaks from crystals containing straight, parallel screw dislocations with precisely known atomic displacements. Though their treatment is mathematically rigorous, it is not possible to determine all of the atomic displacements in a real crystal containing a dislocation network with a diversity of dislocation orientations, shapes, and Burgers vectors. Rago et al.[28–32] have described a phase-invariant model and a mosaic crystal model for dynamical diffraction from metamorphic semiconductor structures in which the ensemble of dislocations is described by characteristic distributions associated with the ensemble average angular and strain broadening of the Bragg profiles by the dislocations. This approach is quite general, but for its application, Rago et al. considered the specific case of Gaussian distributions to calculate the Bragg diffraction profiles from semiconductor structures with nonuniform strain, composition, and dislocation density. Because it had been shown in previous work that curve fitting to dynamical simulations allows the extraction of the strain profiles in semiconductor heterostructures, they presented results for unstrained material, but it is important to note that their general conclusions still hold in the case of strained structures.

7.12.1.1 Phase-Invariant Model

Rago et al. developed a phase-invariant model for the calculation of dynamical x-ray rocking curves from heteroepitaxial structures with nonuniform strain, composition, and dislocation density. It is based on the Takagi–Taupin equation for distorted crystals, but it accounts for the diffuse scattering that arises from mosaicity of strain and angular misorientation associated with the presence of dislocations. It was shown that the diffraction profiles from ZnS_xSe_{1-x} and $In_xGa_{1-x}As/GaAs(001)$ heterostructures are affected strongly by the presence of dislocations, so that the dislocation density can be extracted, as well as strain and composition, through the analysis of experimentally measured rocking curves. Because it would be impossible to completely describe atomic positions of a real structure with its complex network of dislocations, Rago et al. described the dislocations using characteristic statistical distributions associated with the ensemble average angular and strain broadening of the Bragg profiles, which are the two main ways that dislocations distort the lattice. As illustrated in Figure 7.69, some mosaic blocks have angular misorientation (a), while others have a modified spacing of the diffracting planes (b). Both effects serve to modify the angle at which that block diffracts, and this is incorporated into the phase-invariant model by way of the deviation parameter in the dynamical theory of Takagi and Taupin.

This approach is quite general, but for its application Rago et al. considered the specific case of Gaussian distributions for the angular and mosaic spread to calculate the Bragg

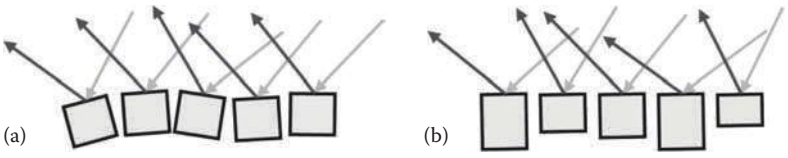


FIGURE 7.69
Effect of (a) angular and (b) interplanar spacing variations of mosaic blocks on diffraction. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

diffraction profiles from semiconductor structures with nonuniform strain, composition, and dislocation density.

7.12.1.1.1 Phase-Invariant Model Theory

In the development of the phase-invariant model, the statistical distribution of angular orientation is defined as $P_\alpha(\theta)$ and the distribution of interplanar spacing variation as $P_\varepsilon(\theta)$. These distributions are then used to modify the deviation parameter given by Equation 7.189, resulting in

$$\eta_n = \left\{ \sqrt{|\gamma_0/\gamma_H|}\, C\Gamma \sqrt{F_{Hn}F_{\bar{H}n}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{P_\alpha(\alpha)P_\varepsilon(\beta)\,d\alpha d\beta}{-(\gamma_0/\gamma_H)(\theta+\alpha-\theta_{Bn}-\beta)\sin(2\theta_{Bn}) - 0.5(1-\gamma_0/\gamma_H)\Gamma F_{0n}} \right\}^{-1}, \tag{7.193}$$

where:

θ_{Bn} is the Bragg angle for the nth sublayer
θ is the actual angle of incidence on the diffracting planes
F_{0n} is the structure factors for the 000 reflection
F_{Hn} is the structure factors for the hkl reflection
$F_{\bar{H}n}$ is the structure factors for the $\bar{h}\bar{k}\bar{l}$ reflection
C is the polarization factor

$$\Gamma = r_e\lambda^2/(\pi V),$$

where:

r_e = 2.818×10^{-5} Å is the classical electron radius
λ is the x-ray wavelength
V is the volume of the unit cell

This effective deviation parameter is discretized and the integrals become a summation for the purposes of computer calculation. The effective deviation parameter given above is calculated and then used in the recursion formulae given earlier for the perfect case, but restated here for clarity, where the scattering amplitude is given by

$$X_n = \eta_n + \sqrt{\eta_n^2 - 1}\,\frac{(S_{1n}+S_{2n})}{S_{1n}-S_{2n}}, \tag{7.194}$$

where

$$S_{1n} = \left(X_{n-1} - \eta_n + \sqrt{\eta_n^2-1}\right)\exp\left(-iT_n\sqrt{\eta_n^2-1}\right) \tag{7.195}$$

and

$$S_{2n} = \left(X_{n-1} - \eta_n + \sqrt{\eta_n^2-1}\right)\exp\left(iT_n\sqrt{\eta_n^2-1}\right). \tag{7.196}$$

The scattering amplitude is first calculated for the substrate, and then it is found recursively for each sublayer until the surface of the structure being modeled has been reached. Though the approach could be applied to arbitrary distributions, Rago et al. considered the specific case of Gaussian distributions for the angle-scale distributions associated with

the variations of orientation and interplanar spacing of the mosaic crystal. In 1953, Gay et al.[120] showed for metals that the distribution of angular variation of mosaic blocks may be modeled as a Gaussian distribution, and in 1994, Ayers later adapted this work for semiconductors,[121] giving the distribution for the angular orientation to be

$$P_\alpha(\theta) = \left(1/\left(\sigma_\alpha\sqrt{2\pi}\right)\right)\exp\left(-\theta^2/\left(2\sigma_\alpha^2\right)\right), \tag{7.197}$$

where the standard deviation is

$$\sigma_\alpha = b\sqrt{\pi D}/2, \tag{7.198}$$

where:

D is the dislocation density
b is the Burgers vector for the dislocations

On the basis of the mosaic block model, it was also shown for metals in 1961 by Hordon and Averbach,[121] and later adapted for semiconductors by Ayers,[122] that for a symmetric reflection from a (001) semiconductor heterostructure with 60° dislocations, the scale distribution $P_\varepsilon(\theta)$ may be modeled as a Gaussian distribution given by

$$P_\varepsilon(\theta) = \left(1/\left(\sigma_\varepsilon\sqrt{2\pi}\right)\right)\exp\left(-\theta^2/\left(2\sigma_\varepsilon^2\right)\right), \tag{7.199}$$

for which the standard deviation is

$$\sigma_\varepsilon = 0.127b\sqrt{D\left|\ln\left(2\times10^{-7}\,\text{cm}\sqrt{D}\right)\right|}\tan\theta_B. \tag{7.200}$$

Determination of the in-plane lattice constant and strain requires XRD measurement of both symmetric and asymmetric reflections. Depicted in Figure 7.70 is a symmetric reflection, while Figure 7.71 illustrates an asymmetric reflection for which the angle between the diffracting planes and the surface is ϕ.

For an asymmetric reflection, Equation 7.193 becomes

$$\eta_n = \left\{\sqrt{\left|\gamma_0/\gamma_H\right|}C\Gamma\sqrt{F_{Hn}F_{\overline{H}n}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{P_\alpha(\alpha)P_\varepsilon(\beta)\,d\alpha d\beta}{-\left(\gamma_0/\gamma_H\right)\left(\theta+\alpha-\theta_{Bn}-\beta+\Delta\phi_{tet}\right)\sin\left(2\theta_{Bn}\right)-0.5\left(1-\gamma_0/\gamma_H\right)\Gamma F_{0n}}\right\}, \tag{7.201}$$

where $\Delta\phi_{tet}$ is the change in ϕ due to tetragonal distortion.

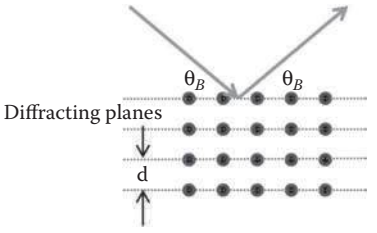


FIGURE 7.70
Symmetric reflection. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

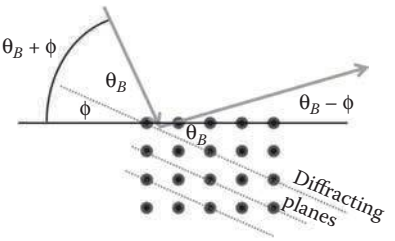


FIGURE 7.71
Asymmetric reflection. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

This framework applies to any *hkl* reflection from a (001)-oriented diamond or zinc blende crystal. It is therefore applicable to III-V/Si(001), II-VI/Si(001), SiGe/Si(001), and II-VI/III-V (001) heterostructures with any arbitrary laminar configuration, and may easily be extended to other crystal orientations through the modification of ϕ and *b*.

7.12.1.1.2 Phase-Invariant Model Results

Rago et al. applied the modified dynamical model for Bragg diffraction from nonuniform semiconductor structures as outlined in the previous section to calculate diffraction profiles of various reflections for $In_xGa_{1-x}As/GaAs(001)$ heterostructures. In their work, the x-ray source wavelength used corresponds to the Cu $k\alpha_1$ line ($\lambda = 0.1540594$ nm) and all rocking curves were convolved with a Gaussian instrumental broadening function with a 20 arcsec FWHM, which corresponds to a standard deviation for the Gaussian of 8.5 arcsec.

Figures 7.72 through 7.74 show the 002, 004, and 006 profiles, respectively, for a 1.0 μm uniform layer of $In_{0.05}Ga_{0.95}As$ with a GaAs(001) substrate and various values of assumed threading dislocation density *D*. Dynamical interactions with the diffracted intensity from the substrate yield an asymmetric $In_{0.05}Ga_{0.95}As$ Bragg peak profile, and this is most pronounced in the case of the 002 reflection. For all three reflections, the FWHM of the $In_{0.05}Ga_{0.95}As$ peak increases monotonically as *D* is increased from 0 to 5×10^7 cm^{-2}, showing that the dislocation density may be determined from the FWHM.

The convolution model[121] provides a simple hand calculation for estimating dislocation density from the width of the Bragg peak for a uniform layer, and it is based on the same mosaic block model as the phase-invariant model for dynamical diffraction. To compare the detailed dynamical simulations with the convolution model, Figure 7.75 gives the rocking curve FWHM for the 002, 004, and 006 Bragg reflections as a function of the dislocation density *D*. Also shown in Figure 7.75 is the rocking curve width calculated using the convolution model,[121] for which

$$\beta^2 = \beta_0^2 + \beta_\alpha{}^2 + \beta_\varepsilon{}^2 = \beta_0^2 + 2\pi b^2 D \ln 2 + 0.16 b^2 D \left| \ln\left(2\times 10^{-7}\sqrt{D}\right)\right| \tan^2\theta_B, \tag{7.202}$$

where:

β_0 is the intrinsic rocking curve width, determined using a dynamical simulation for a perfect crystal with the same thickness

β_α and β_ε are the FWHMs of the angle-scale distributions associated with angular and strain mosaic spread, respectively.

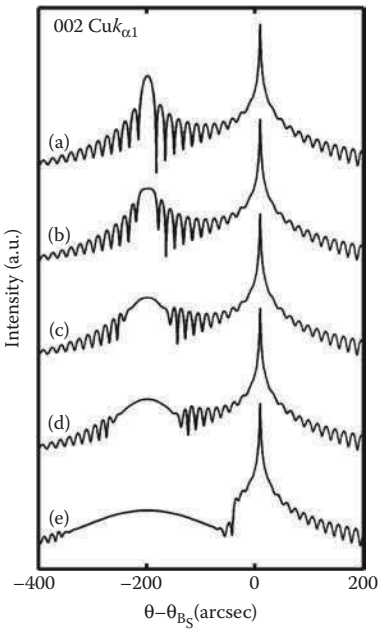


FIGURE 7.72
002 rocking curves for uniform 1.0 μm $In_{0.05}Ga_{0.95}As$/As(001) with D equal to (a) 0 cm^{-2}, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, and (e) 5×10^7 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

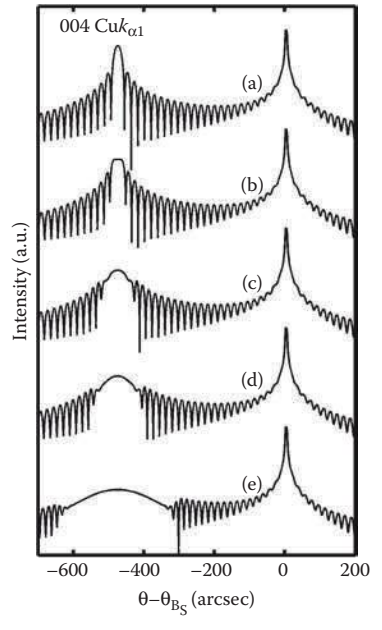


FIGURE 7.73
004 rocking curves for uniform 1.0 μm $In_{0.05}Ga_{0.95}As$/As(001) with D equal to (a) 0 cm^{-2}, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, and (e) 5×10^7 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

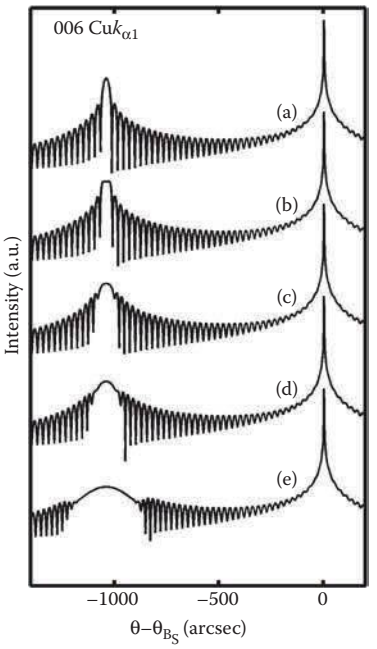


FIGURE 7.74
006 Rocking curves for uniform 1.0 μm $In_{0.05}Ga_{0.95}As$/As(001) with D equal to (a) 0 cm^{-2}, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, and (e) 5×10^7 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

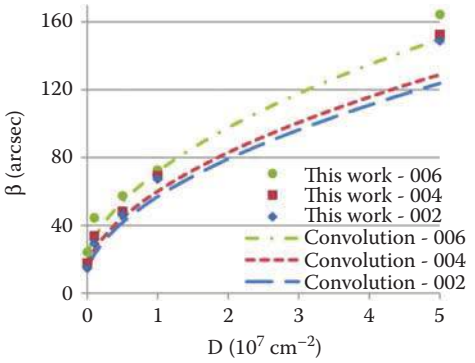


FIGURE 7.75
Comparison of Bragg peak width with that predicted by the convolution model.[121] (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

There is good agreement between the dynamical calculations and the convolution model even at the lower dislocation densities, which are near the lower limit for the use of x-ray characterization.

Rago et al. considered seven different *hkl* reflections from a 1.0 μm uniform layer of $In_{0.05}Ga_{0.95}As$/GaAs(001) with constant $D = 5 \times 10^6$ cm^{-2}. Shown in Figure 7.76 are the 002, 004, 044, 135, 006, 026, and 335 profiles, which are in order of increasing Bragg angle.

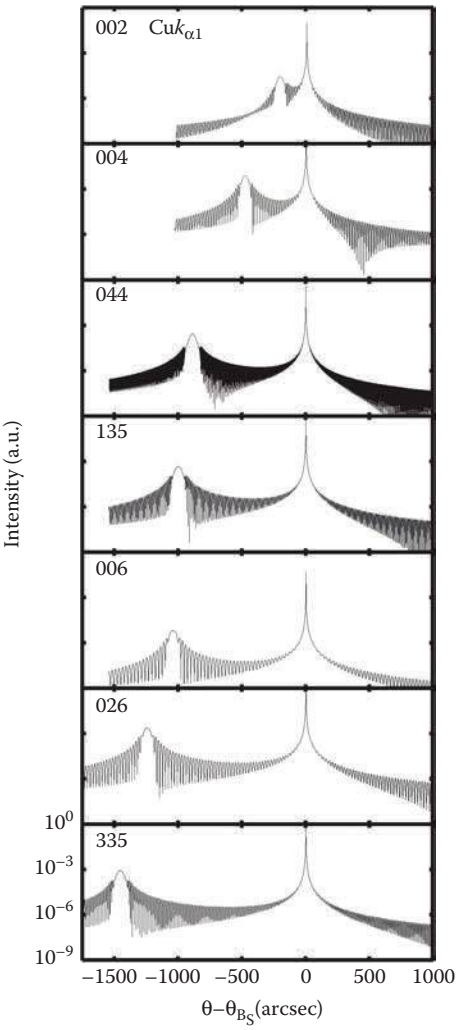


FIGURE 7.76
Rocking curves for various reflections from 1.0 μm $In_{0.05}Ga_{0.95}As/GaAs(001)$ with $D = 5 \times 10^6$ cm^{-2}. Both the intensity and angular scales are the same in all plots. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

In the simplest form, the convolution model is applied by neglecting the intrinsic width of the rocking curve width β_o, including only the widths due to angular (β_α) and strain (β_ε) broadening. This is appropriate for uniform samples with high, uniform dislocation densities, for which $\beta_\alpha^2 + \beta_\varepsilon^2 \gg \beta_o^2$ and yields a linear relationship between the square width β^2 and $\tan^2 \theta_B$, which is given by

$$\beta^2 = \beta_\alpha^2 + \beta_\varepsilon^2 = 2\pi b^2 D \ln 2 + 0.16 b^2 D \left| \ln\left(2\times 10^{-7}\sqrt{D}\right)\right| \tan^2 \theta_B. \tag{7.203}$$

The approximate analysis given by Equation 7.203 (excluding the intrinsic rocking curve width) is plotted in Figure 7.77, as well as the widths from the various reflections for the

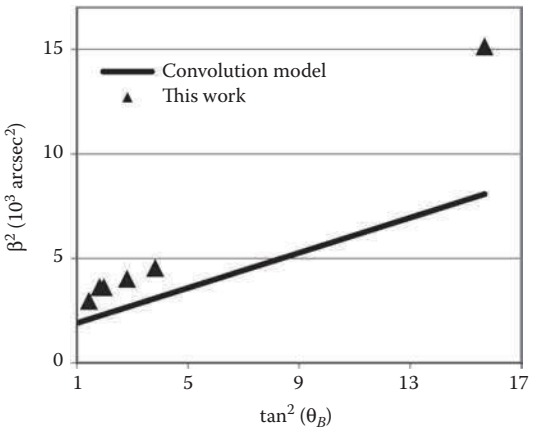


FIGURE 7.77
Comparison of squared width vs. $\tan^2(\theta_B)$ for various reflections with the convolution model[121] from a 1.0 μm $In_{0.05}Ga_{0.95}As/GaAs(001)$ structure with $D = 5 \times 10^6$ cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

$D = 5 \times 10^6$ cm^{-2} sample calculated by the phase-invariant model, versus $\tan^2(\theta_B)$ for various reflections.

It is seen in Figure 7.77 that the convolution model greatly underestimates the width when the intrinsic width is excluded for the dislocation density considered. When including the intrinsic width of the epilayer, the convolution model agrees to within ~20%. Rago et al. also showed that an important advantage of their model is the ability to analyze nonuniform structures for which the convolution model cannot be applied.

To further demonstrate this, 044 diffraction profiles were calculated for 1.0 μm $In_{0.07}Ga_{0.93}As/GaAs(001)$ structures with the epilayer containing linear dislocation density profiles. In these structures, D varied linearly from D_1 at the substrate interface to D_2 at the surface. Figure 7.78 shows the calculated rocking curves for the structures with (D_1, D_2) equal to (2×10^7, 0), (1.5×10^7, 0.5×10^7), (1.2×10^7, 0.8×10^7), and (1×10^7, 1×10^7) cm^{-2}. All four samples have the same average defect density of 10^7 cm^{-2}, yet the rocking curve peak intensity, width, and shape are all influenced by the dislocation distribution. Rago et al. concluded on the basis of this observation that, in principle, it should be possible to extract the dislocation density distribution by curve fitting to a measured HRXRD profile.

This raises the question of whether this would also hold for structures having nonuniform composition. To address this, Rago et al. calculated diffraction profiles for multilayer structures comprising 0.3 μm $In_{0.06}Ga_{0.94}As$/0.3 μm $In_{0.07}Ga_{0.93}As$/0.3 μm $In_{0.08}Ga_{0.92}As$/0.3 μm $In_{0.09}Ga_{0.91}As/GaAs(001)$, with linear depth profiles of dislocation throughout the entire epitaxial structure, and the results are given in Figure 7.79 for linear dislocation profiles (D_1, D_2) of (2×10^7, 0), (1.5×10^7, 0.5×10^7), (1.2×10^7, 0.8×10^7), and (1×10^7, 1×10^7) cm^{-2}. Here, the value of D_1 corresponds to the defect density at the substrate interface and D_2 corresponds to that at the surface. Also shown in the figure is the diffraction profile for the dislocation-free structure with $D_1 = D_2 = 0$. They found that the rocking curve varies with the dislocation distribution even though the average dislocation density was fixed in all structures (excluding the $D = 0$ case). From this, it was concluded that the rocking curve is indeed sensitive to the depth profile of the dislocation density and not just the average density.

Rago et al. also considered other material systems, including ZnSSe/GaAs(001). Rocking curves were calculated for a $ZnS_xSe_{1-x}/GaAs(001)$ multilayer structure, comprised of

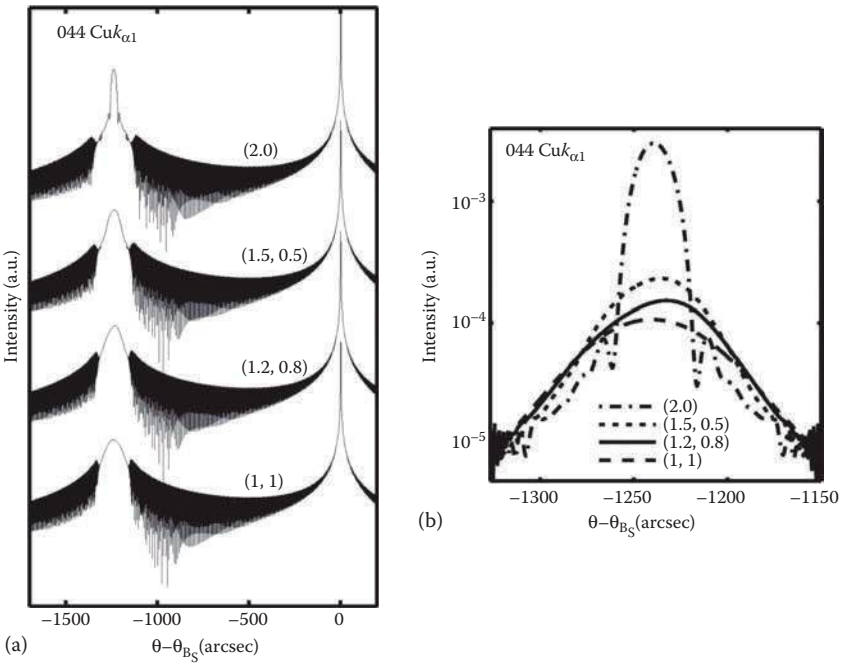


FIGURE 7.78
044 Rocking curves for a 1 μm $In_{0.07}Ga_{0.93}As/GaAs(001)$ structure with linear dislocation density profiles given by (D_1, D_2) in 10^7 cm^{-2}, for (a) both the substrate and epitaxial Bragg peaks, and (b) detail of the epilayer peak. D_1 is the dislocation density at the substrate interface and D_2 is the dislocation density at the surface.

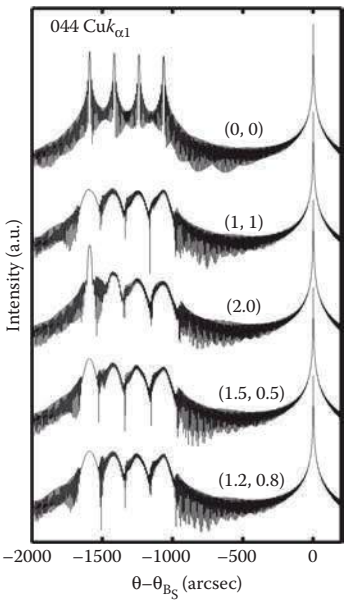


FIGURE 7.79
044 Rocking curves for $In_xGa_{1-x}As/GaAs(001)$ multilayer structure with linear dislocation density depth profiles (D_1, D_2) in 10^7 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

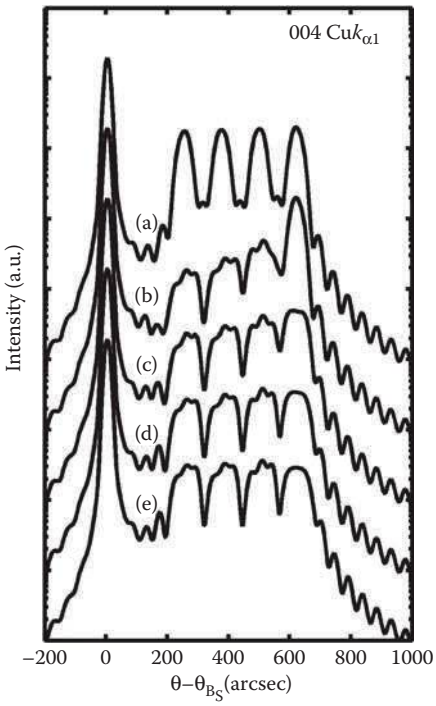


FIGURE 7.80
004 Rocking curves for ZnS_ySe_{1-y}/GaAs(001) multilayers with various linear distributions of dislocation density in which (D_1, D_2) are equal to (a) (0, 0), (b) (2×10^7, 0), (c) (1.1×10^7, 0.5×10^7), (d) (1.2×10^7, 0.8×10^7), and (e) (1×10^7, 1×10^7) cm^{-2}. Reprinted from P. B. Rago et al., *J. Electron. Mater.*, 36, [2012]. With permission. Copyright 2012, Minerals, Metals and Materials Society.)

0.4 μm $ZnS_{0.16}Se_{0.84}$/0.4 μm $ZnS_{0.14}Se_{0.86}$/0.4 μm $ZnS_{0.12}Se_{0.88}$/0.4 μm $ZnS_{0.10}Se_{0.90}$ top layers on a GaAs(001) substrate. Rocking curves for these structures, all identical in their compositional profile, were simulated using the phase-invariant model with linear dislocation density depth profiles (D_1, D_2) equal to (a) (0, 0), (b) (2×10^7, 0), (c) (1.5×10^7, 0.5×10^7), (d) (1.2×10^7, 0.8×10^7), and (e) (1×10^7, 1×10^7) cm^{-2}, where D_1 is the dislocation density at the substrate interface and D_2 is the dislocation density at the surface. The 004 diffraction profiles for this structure are given in Figure 7.80. Similar to the case of the $In_xGa_{1-x}As$/GaAs(001) material system, the rocking curve varies with the depth profile of dislocation density even though the average dislocation density is fixed at 1×10^7 cm^{-2}, further confirming the sensitivity of the rocking curve to the depth profile of the dislocation density.

Rago et al. also considered superlattice structures, which are of interest for device applications, for control of defect density, and for growth characterization experiments. They showed that $In_xGa_{1-x}As$ superlattices on GaAs(001) substrates exhibit rather complex diffraction behavior with nonuniform dislocation distributions, even using simple linear profiles for D, and concluded that more work would be necessary to clarify the characterization of the dislocation density in the general case of mutlilayers and superlattices. Figure 7.81 shows 044 rocking curves for a 10-period 10-nm $In_{0.03}Ga_{0.97}As$/10-nm $In_{0.04}Ga_{0.96}As$ superlattice on GaAs(001) with four different linear profiles of dislocation density in which (D_1,D_2) are equal to (a) (0, 0), (b) (1×10^7, 1×10^7), (c) (2×10^7, 0), (d) (1.5×10^7, 0.5×10^7), and (e) (1.2×10^7, 0.8×10^7) cm^{-2}.

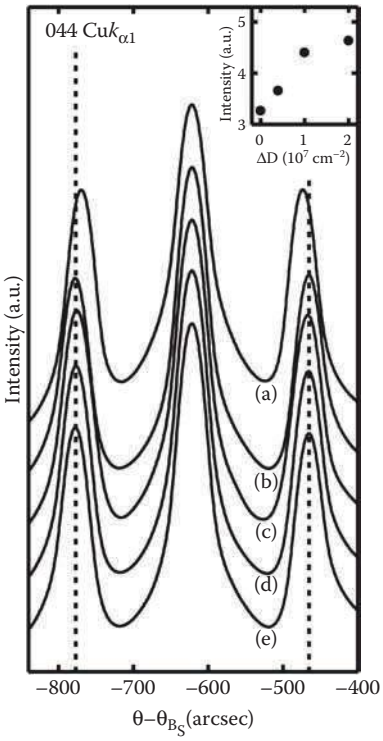


FIGURE 7.81
044 Diffraction profiles for a 10-period 10 nm $In_{0.03}Ga_{0.97}As$/10 nm $In_{0.04}Ga_{0.96}As$ superlattice on GaAs(001) with different linear distributions of dislocation density in which (D_1, D_2) are equal to (a) (0, 0), (b) (1×10^7, 1×10^7), (c) (2×10^7, 0), (d) (1.5×10^7, 0.5×10^7), and (e) (1.2×10^7, 0.8×10^7) cm^{-2}. Inset shows average peak intensity I of the +1 and −1 superlattice peaks as a function of $\Delta D = D_2 - D_1$; reference angle lines show variation of the angle of superlattice peaks. (Reprinted from P. B. Rago and J. E. Ayers, *J. Electron. Mater.*, 42, 2450 [2013]. With permission. Copyright 2013, Minerals, Metals, and Materials Society.)

7.12.1.2 Dynamical Modeling of Asymmetrical Dislocation Densities

The dynamical theory of Bragg XRD has been extended[30] to the modeling of asymmetric dislocation densities, which are described in detail in Section 7.11.4. In a zinc blende semiconductor, there are eight active slip systems. The two distinct types of dislocations that exist in these slip systems are distinguished by their misfit segments, which are oriented along either the [110] or [1–10] direction. The two threading dislocation densities can be estimated by measuring the x-ray rocking curve width as a function of incident beam azimuth. However, the crystallographic tilt that may be present in the epilayer also has a measurable and conflicting effect on the rocking curve with respect to sample azimuth. First, within a full 360° azimuthal rotation the layer's Bragg peak position varies by nominally twice the layer's absolute tilt. Second, the epilayer tilt affects the rocking curve width, though we will see that this effect is only of the second order, so that the layer misorientation with respect to the substrate need not be considered for the purpose of determining the two dislocation populations.

First, to account for crystallographic tilting in the epitaxial layers, the deviation parameter of the phase-invariant model (Equation 7.176) is modified to

$$\eta_n = \left\{ \sqrt{|\gamma_0 / \gamma_H|}\, C\Gamma \sqrt{F_{Hn} F_{\overline{H}n}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P_\alpha(\alpha) P_\varepsilon(\beta)\, d\alpha d\beta}{-(\gamma_0 / \gamma_H)\Omega \sin(2\theta_{Bn}) - 0.5(1 - \gamma_0 / \gamma_H)\Gamma F_{0n}} \right\}^{-1} \tag{7.204}$$

where

$$\Omega = \theta - \theta_{Bn} + \alpha - \beta + \Delta\phi_{\text{tet}} + \Delta\phi \cos(\psi - \psi_0), \tag{7.205}$$

ϕ is the substrate inclination
$\Delta\phi_{\text{tet}}$ is the change in ϕ due to tetragonal distortion
$\Delta\phi$ is the epilayer tilt with respect to the substrate
γ_0/γ_H is the asymmetry factor

The direction cosines become

$$\gamma_0 = \sin\left\{\theta_B + \left[\phi + \Delta\phi_{\text{tet}} + \Delta\phi \cos(\psi - \psi_0)\right]\right\} \tag{7.206}$$

and

$$\gamma_H = -\sin\left\{\theta_B - \left[\phi + \Delta\phi_{\text{tet}} + \Delta\phi \cos(\psi - \psi_0)\right]\right\}. \tag{7.207}$$

The notion that there are two separate types of 60° dislocations in a zinc blende crystal was introduced in 1972 by Abrahams et al.[181] We can classify the eight active slip systems as Type A and Type B, according to their misfit dislocation line vector, with orientation along the [110] and [1–10] directions, respectively. It can be considered that α-type dislocations are associated with Type A slip systems, and β-type dislocations are associated with Type B systems. It should be noted that the two types are not chemically equivalent; for example, in the GaAs material system, α-type dislocations have Ga atoms at their core, while β-type all have As atoms at their core.[180] The dislocation density of the α-type dislocations is defined as D_A, and that of the β-type dislocations as D_B.

Because the two networks of dislocations are orthogonal, their populations can be measured separately by varying the incident x-ray beam azimuth. Yarlagadda et al.[167] derived a quantitive model that estimates the two defect densities, and this work has been incorporated into the more detailed phase-invariant dynamical model.

Yarlagadda obtained the result that the squared width of the rocking curve varies with the azimuth ψ as

$$\beta^2 = 8\pi \ln 2 b^2 \left(D_A \cos^2\psi \cos^2\gamma_A + D_B \sin^2\psi \cos^2\gamma_B\right). \tag{7.208}$$

Since the FWHM is related to the standard deviation of a Gaussian by $\sigma = \beta/(2\sqrt{2\ln 2})$, the angular mosaic spread can be described by

$$\sigma_\alpha = b\sqrt{\pi}\left(D_A \cos^2\psi \cos^2\gamma_A + D_B \sin^2\psi \cos^2\gamma_B\right)^{1/2}. \tag{7.209}$$

For the 004 reflection, this simplifies to

$$\sigma_\alpha = b\sqrt{\pi/2}\left(D_A \cos^2\psi + D_B \sin^2\psi\right)^{1/2}. \tag{7.210}$$

The strain mosaic spread is modified simply by letting $D = D_A + D_B$ in the expression for the standard deviation of strain mosaic spread so that

$$\sigma_\varepsilon = 0.127 b \sqrt{(D_A + D_B)\left|\ln\left(2\times10^{-7}\,\text{cm}\sqrt{D_A + D_B}\right)\right|}\,\tan\theta_B. \tag{7.211}$$

Note that this result does not depend on azimuth, and this may be explained physically as follows. For a crystallite affected by the presence of dislocations in such a way that it is under biaxial strain, but has no angular misorientation, the biaxial strain causes tetragonal distortion, and this alters the spacing of the (*hkl*) diffracting planes. With no angular tilt component with respect to the [*hkl*] direction, the diffraction from the crystallite is equal as the sample is rotated about its azimuth.

This model has been applied by Rago et al. to an asymmetrically dislocated structure consisting of a 1 μm layer of $Hg_{0.5}Cd_{0.5}Te$ on a CdTe(001) substrate. The assumed dislocation densities in the epitaxial layer were equal to $D_A = 5 \times 10^7$ cm^{-2} and $D_B = 10^8$ cm^{-2}. Depicted in Figure 7.82 are the 004 rocking curves for azimuths of $\psi = 0°$, 45°, and 90° of this structure. It is seen in Figure 7.83 that the square of the epilayer peak width follows the expected dependence on ψ of Equation 7.208.

Rago et al.[30] compared experimental measurements made by Yarlagadda et al. in a 2008 paper[167] with diffraction profiles calculated using the dynamical theory modified for asymmetric dislocations. In the 2008 work, Yarlagadda et al. grew a 380 nm ZnSe layer on GaAs(001), and measured x-ray rocking curves at 18 azimuthal angles. From this study, they concluded that the dislocation density of the α-type dislocations was $D_A = 2 \times 10^8$ cm^{-2} and of the β-type dislocations was $D_B = 2.5 \times 10^8$ cm^{-2}. Rocking curves were calculated for this structure, adjusting the dislocation densities until a good match was achieved for the square width. The values that were needed to achieve the closest match were $D_A = 1.6 \times 10^8$ cm^{-2} and $D_B = 2.0 \times 10^8$ cm^{-2}, and the resulting characteristic is given in Figure 7.84, along with the experimental measurements. The dislocation densities

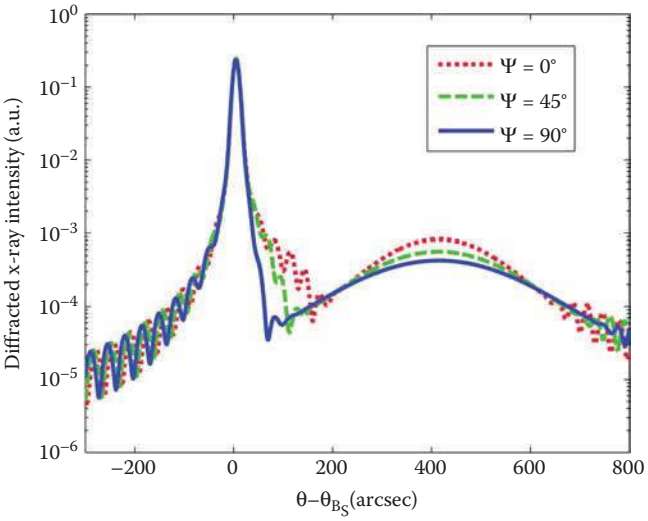


FIGURE 7.82
004 Rocking curves for HgCdTe/CdTe with $D_A = 5 \times 10^7$ cm^{-2} and $D_B = 10^8$ cm^{-2} at $\psi = 0°$, 45°, and 90°. (Reprinted from P. B. Rago et al., *J. Electron. Mater.*, 42, 3066 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

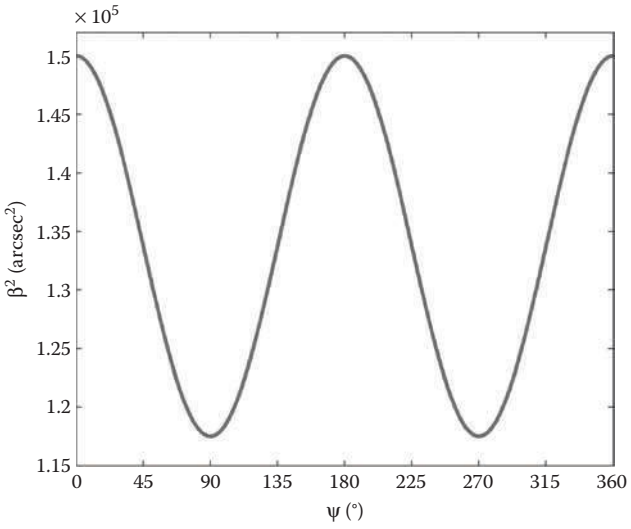


FIGURE 7.83
Square of the FWHM vs. azimuth for the 004 reflections from $Hg_{0.5}Cd_{0.5}Te/CdTe(001)$ with $D_A = 5 \times 10^7$ cm^{-2} and $D_B = 10^8$ cm^{-2}. (Reprinted from P. B. Rago et al., *J. Electron. Mater.*, 42, 3066 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

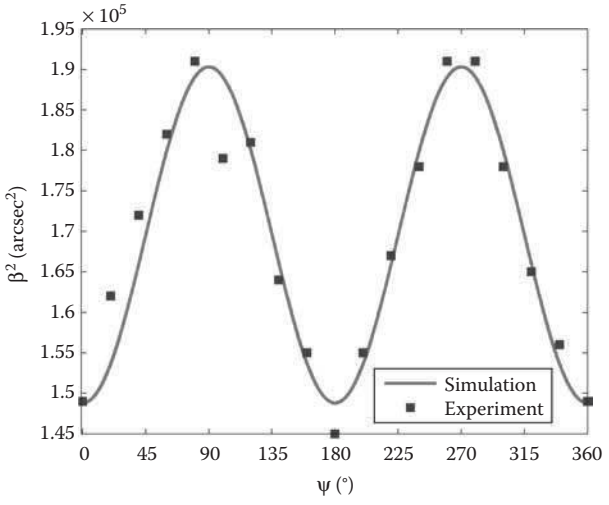


FIGURE 7.84
Simulated and experimental square width vs. azimuth for a 380 nm ZnSe/GaAs(001) Structure. (Reprinted from P. B. Rago et al., *J. Electron. Mater.*, 42, 3066 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

so obtained were roughly 20% less than those found by Yarlagadda et al., but the lower values were expected to be more accurate because they involved dynamical calculations. Nevertheless, the simple hand analysis also provides fair accuracy.

Therefore, the two dislocation populations may be determined with fair accuracy using either dynamical simulations or hand analysis of the width squared versus azimuth characteristic.

7.12.1.3 Mosaic Crystal Model

The phase-invariant model for dynamical diffraction presented in the previous sections is broadly applicable to semiconductor heterostructures and extends the x-ray analysis to determination of the depth profile of the dislocation density, as well as the strain and composition. Although it is applicable to many metamorphic semiconductor heterostructures, it does not account for phase differences between the crystallites of the mosaic. This leads to incorrect predictions of certain dynamic effects, including extinction and Pendellosung interference fringes for some cases. The model therefore lacks accuracy in certain structures, such as those containing either closely lattice matched layers or layers with very high threading dislocation densities. The inaccuracy in some technically important cases of metamorphic structures was the motivation for the development of a refined mosaic crystal model by Rago et al., and this refined model is more generally applicable to semiconductor device structures. Specifically, it is more appropriate for step-graded metamorphic semiconductor structures, as well as any other type of semiconductor heterostructure containing closely lattice matched layers or dislocation densities higher than 10^8 cm^{-2}. Additionally, it is significantly more computationally efficient, so rocking curves for complex structures with many sublayers or high dislocation densities can be computed in much less time. The following sections describe the mosaic crystal model and its application to step-graded $In_xGa_{1-x}As/GaAs(001)$ and $In_xAl_{1-x}As/GaAs(001)$ device structures.

7.12.1.3.1 Mosaic Crystal Model Theory

The starting point for the mosaic crystal model is the assumption that dislocations give rise to angular variations, as well as variations in the interplanar spacing, and that these effects may be modeled by the distributions $P_\alpha(\theta)$ and $P_\varepsilon(\theta)$, respectively. The general approach for the phase-invariant dynamical diffraction model (PIDDM) and mosaic crystal dynamical diffraction model (MCDDM) are as follows: in the phase-invariant model, the angular and mosaic spread $P_\alpha(\theta)$ and $P_\varepsilon(\theta)$ are utilized to determine an effective deviation parameter that is computed through the double integration with respect to these distributions, and the rocking curve is computed with this effective deviation parameter. In contrast, application of the mosaic crystal model involves calculation of the scattering amplitude for each of an array of mosaic blocks, keyed to the two distributions, and the intensities are weighted and summed.

For this purpose, the structure is considered to comprise a mosaic of $N_\alpha \times N_\beta$ crystallites in α and β space. The dynamic scattering amplitude is computed for each of these crystallites iteratively. For the nth lamina of one such crystallite, the deviation parameter is

$$\eta_{nij} = \frac{-(\gamma_0/\gamma_H)(\theta - \theta_{Bn} + \alpha_i - \beta_j)\sin(2\theta_{Bn}) - 0.5(1-\gamma_0/\gamma_H)\Gamma F_{0n}}{\sqrt{|\gamma_0/\gamma_H|}C\Gamma\sqrt{F_{Hn}F_{\bar{H}n}}}, \tag{7.212}$$

where $\alpha_i = -N_\sigma\sigma_\alpha + iN_\alpha\sigma_\alpha/2N_\sigma$ and $\beta_j = -N_\sigma\sigma_\beta + jN_\beta\sigma_\beta/2N_\sigma$, and where N_σ is the number of standard deviations used in the two distributions and i and j are integers. The scattering amplitudes are calculated iteratively by

$$X_{nij} = \eta_{nij} + \sqrt{\eta_{nij}^2 - 1}\,\frac{(S_{1nij} + S_{2nij})}{S_{1nij} - S_{2nij}}, \tag{7.213}$$

where

$$S_{1nij} = \left(X_{nij-1} - \eta_{nij} + \sqrt{\eta_{nij}^2 - 1} \right) \exp\left(-iT_n \sqrt{\eta_{nij}^2 - 1} \right) \tag{7.214}$$

and

$$S_{2nij} = \left(X_{nij-1} - \eta_{nij} + \sqrt{\eta_{nij}^2 - 1} \right) \exp\left(iT_n \sqrt{\eta_{nij}^2 - 1} \right), \tag{7.215}$$

where η_n is the deviation parameter and the thickness parameter is

$$T_n = h_n \frac{\pi \Gamma \sqrt{F_{Hn} F_{\bar{H}n}}}{\lambda \sqrt{|\gamma_0 \gamma_H|}}, \tag{7.216}$$

where h_n is the thickness and F_{0n}, F_{Hn}, and $F_{\bar{H}n}$ are the 000, *hkl*, and $\bar{h}\,\bar{k}\,\bar{l}$ structure factors for the *n*th sublayer. The XRD profile is calculated by adding the weighted intensity contributions from the $N_\alpha \times N_\beta$ crystallites in α and β space with

$$I = \sum_i \sum_j \left| X_{Nij} \right|^2 \cdot W_{\alpha i} \cdot W_{\varepsilon j}, \tag{7.217}$$

where the weighting functions are given by

$$W_{\alpha i} = \exp\left(\frac{-\alpha_i^2}{2\sigma_\alpha^2} \right), \tag{7.218}$$

and

$$W_{\varepsilon i} = \exp\left(\frac{-\beta_i^2}{2\sigma_\varepsilon^2} \right), \tag{7.219}$$

and the standard deviations may be related to the dislocation density by

$$\sigma_\alpha = b\sqrt{\pi D/2}, \tag{7.220}$$

and

$$\sigma_\varepsilon = 0.127 b \sqrt{D \ln\left(2 \times 10^{-7}\ \text{cm} \sqrt{D} \right) \tan\theta_B}\,. \tag{7.221}$$

This approach accounts for the incoherency of the phase within crystallites of the mosaic, and therefore is a more physically accurate model of diffraction within a distorted crystal with mosaic block angular misorientations and interplanar variations. The mosaic crystal model may be applied to semiconductor heterostructures with arbitrary profiles of composition and dislocation density.

7.12.1.3.2 Mosaic Crystal Model Results

Rago et al. applied the mosaic crystal model to calculate XRD profiles for step-graded $In_xGa_{1-x}As/GaAs(001)$ structures. The x-ray wavelength corresponded to the Cu $k\alpha_1$ radiation line, $\lambda = 0.1540594$ nm. The layers of the step-graded structure were assumed to have uniform dislocation density. Rago et al. chose $N_\sigma = 3$ because it was found that a larger number of standard deviations did not provide an increase in accuracy. The lower limit on the number of crystallites used was set to $N_\alpha = 21$ and $N_\beta = 21$, resulting in a model involving 441 crystallites. It was noted that the required number of crystallites for accurate modeling depends on the maximum dislocation density in the structure, and therefore the best practice is to increase the number of crystallites until the diffraction profile exhibits no change. For structures with an assumed dislocation density of $D = 10^8$ cm^{-2}, they found that $N_\alpha = N_\beta = 31$ was necessary.

Figure 7.85 shows 004 rocking curves for 1.0 μm step-graded layers of $In_xGa_{1-x}As/GaAs(001)$, with the composition graded from 0% to 50% indium in five steps, and for five different dislocation densities. Although five diffraction peaks are easily resolved in the dislocation-free structure, these diffraction peaks broaden and begin to overlap as the dislocation density increases. In addition, the Pendellosung begin to extinguish when the dislocation density exceeds 10^7 cm^{-2}.

Figure 7.86 shows the 004 rocking curves for a similar structure with 10 steps. The 10 diffraction peaks associated with the compositions in the step-graded structure become increasingly difficult to resolve with higher dislocation density, and here too there is suppression of the Pendellosung with increasing dislocation density.

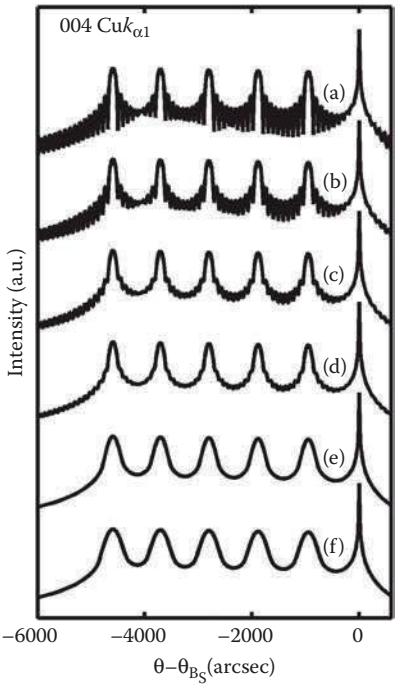


FIGURE 7.85
004 Diffraction profile for step-graded InGaAs/GaAs (five steps) with dislocation density of (a) 0, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, (e) 5×10^7 cm^{-2}, and (f) 10^8 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, in press [2016]. With permission. Copyright 2016, World Scientific.)

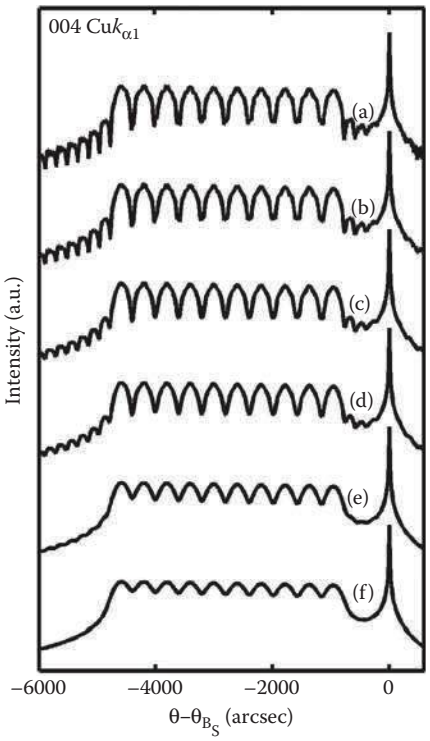


FIGURE 7.86
004 Diffraction profile for step-graded InGaAs/GaAs(10 steps) with dislocation density of (a) 0, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, (e) 5×10^7 cm^{-2}, and (f) 10^8 cm^{-2}. (Reprinted from P. B. Rago and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, in press [2016]. With permission. Copyright 2016, World Scientific.)

Another practical case is that of nonuniform dislocation density. Figure 7.87 shows the 004 rocking curves for 1.0 μm step-graded layers of $In_xGa_{1-x}As/GaAs(001)$ with the composition graded from 0% to 50% indium in five steps, and with the dislocation density decreased stepwise, in a linear fashion, in the step-graded layers. The dislocation density in the first layer (the layer contacting the substrate) was assumed to be 10^8 cm^{-2} in all cases, and the dislocation density in the following layers decreased linearly to a top layer value of (a) 10^8, (b) 5×10^7, (c) 10^7, (d) 5×10^6, (e) 10^6, and (f) 0 cm^{-2}, as indicated in Figure 7.87. Differences are noted in the diffraction from each step in the graded buffer, and as in the uniform case, they observed Pendellosung in layers containing the lower dislocation densities, and the fringes become extinguished as D increases. These results showed that the XRD profile curve is sensitive to the profile of the dislocation density, as well as the average value, as was found for the approximate phase-invariant model.

Finally, Rago et al. considered a device heterostructure grown by Jiang et al. in a 2008 work.[182] The structure comprises a GaAs substrate, a 900 nm overshoot step-graded $In_xAl_{1-x}As$ buffer with nine steps from x = 0.05 to x = 0.85, and a 1 μm $In_{0.75}Al_{0.25}As$ top layer (see the lattice profile for this structure in Figure 7.88). For the purpose of simulation, Rago et al. assumed the structure to be completely relaxed, as suggested by Jiang et al. Shown in Figure 7.89 is the experimental 004 rocking curve, along with that calculated using the mosaic crystal model, in which the top layer was assumed to have a uniform dislocation

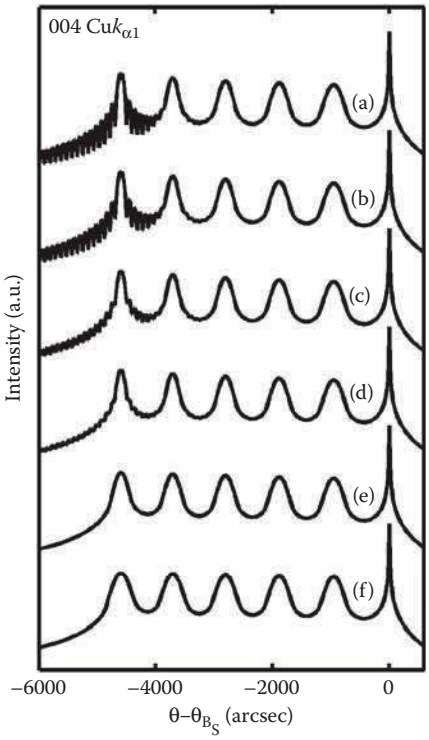


FIGURE 7.87
004 Diffraction profile for step-graded InGaAs/GaAs (five steps) with $D = 10^8$ cm^{-2} at the interface varied stepwise linearly to (a) 0, (b) 10^6 cm^{-2}, (c) 5×10^6 cm^{-2}, (d) 10^7 cm^{-2}, (e) 5×10^7 cm^{-2}, and (f) 10^8 cm^{-2} at the top layer. (Case (f) represents a uniform dislocation density.) (Reprinted from P. B. Rago and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, in press [2016]. With permission. Copyright 2016, World Scientific.)

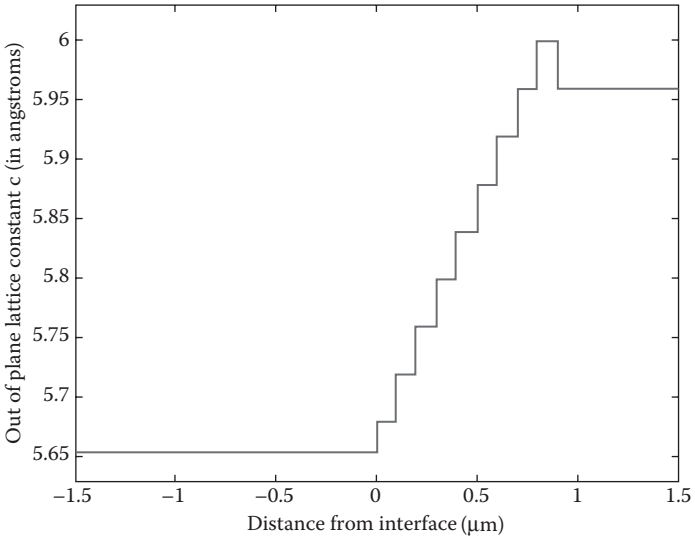


FIGURE 7.88
Out-of-plane lattice constant profile for the $In_xAl_{1-x}As/GaAs(001)$ structure grown by Jiang et al.[181]

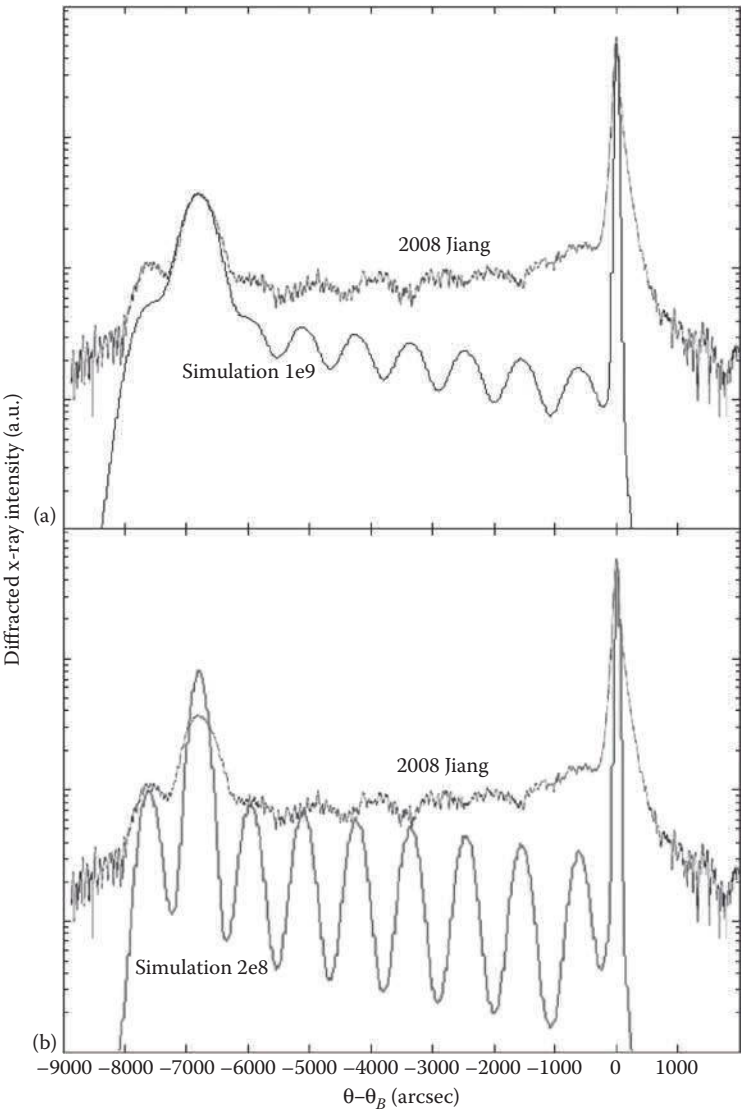


FIGURE 7.89
Comparison of an experimentally measured 004 diffraction profile[29] with diffraction profiles calculated with the mosaic crystal model. The dislocation density in the top layer was assumed to be (a) $D = 10^9$ cm^{-2} and (b) $D = 2 \times 10^8$ cm^{-2}, and the dislocation density in the graded buffer was assumed to be 10^9 cm^{-2} in both cases. (Reprinted from P. B. Rago and J. E. Ayers, *Intl. J. High Speed Electron. Syst*., in press [2016]. With permission. Copyright 2016, World Scientific.)

density of (a) 10^9 cm^{-2} and (b) 2×10^8 cm^{-2}, and the entire step-graded buffer was calculated with $D = 10^9$ cm^{-2} in both cases.

In the case where the dislocation density was assumed to be 10^9 cm^{-2}, the predicted profile of the 1 μm $In_{0.75}Al_{0.25}As$ layer exhibited a width (~600 arcsec) and a normalized intensity (normalized by dividing by the substrate peak intensity, ~0.06) that were in good agreement with the experimental profile shown in Figure 7.89a. On the other hand, if the top $In_{0.75}Al_{0.25}As$ layer was assumed to have a lower dislocation density, then the peak width

was overestimated while the intensity was underestimated. This is seen in Figure 7.89b, for which the top layer was assumed to have a dislocation density of $D = 2 \times 10^8$ cm^{-2}. The resulting width and normalized intensity were 280 in. and 0.15, respectively. These observations suggest that the actual dislocation density in the top uniform layer is approximately 10^9 cm^{-2}.

In the graded buffer, neither calculated diffraction profile achieved good agreement with the measurement. The experimental profile exhibits shifts in the peak positions, and there are several possible explanations for this. Perhaps the most obvious explanation for this is the presence of residual strain in the crystal sample, though Jiang et al. concluded the structure was completely relaxed. The peak shifts are not uniform; that is, each step in the buffer is shifted by a different amount, and this could indicate that the depth profile of residual strain is also nonuniform. If this is the correct explanation, the layers closer to the substrate appear to include tensile in-plane strain; however, the mismatch strain in the structure is compressive. Another possible explanation is the tolerance on compositional control during growth. If the layers closer to the substrate were indium deficient compared with their target composition, then this could cause the observed peak shift. Finally, it is also possible that crystallographic tilting of the layers caused the peak shift. Tilted epitaxial growth is known to occur, with several mechanisms of action proposed.[106] It is possible that the steps in the graded buffer contain different tilts, shifting their Bragg peaks accordingly. To separate these three effects—residual strain, compositional tolerance, and epitaxial tilt—one could measure asymmetric *hkl* reflections and vary the azimuth. Another disagreement between the simulation and experiment is the difference in peak intensities of the peaks associated with the graded buffer. The noise floor of the experimental data could explain this in part, but not entirely. Also, unexpected variations in the growth rate would lead to different thicknesses of the steps, and this could help explain the differences observed in intensity.

Up to this point in time, very little work has been done by way of comparing the mosaic crystal model with experimental structures, so that much more work is necessary to clarify the issues identified above. Nonetheless, the groundwork described above demonstrates that, in principle at least, dynamical simulations should enable the nondestructive depth profiling of the dislocation density in these structures.

7.13 Characterization of the Growth Mode

It is important to be able to characterize the growth mode for the development of heteroepitaxial devices. In many applications, two-dimensional (either Frank–van der Merwe or step-flow) growth is necessary for the attainment of flat interfaces in devices. On the other hand, quantum dot devices require a Volmer–Weber or Stranski–Krastanov growth mode. In all of these cases, however, it is valuable to be able to characterize the growth mode so the behavior can be understood and controlled.

In the case of MBE, the growth mode can be studied *in situ* using RHEED. A streaky RHEED pattern corresponds to an atomically smooth surface, whereas a spotty pattern indicates surface roughening, which is an indication of Volmer–Weber of Stranski–Krastanov growth. If the growth mode is Stranski–Krastanov, layer-by-layer growth of the wetting layer will be followed by islanding. In this case, RHEED oscillations will be seen during the growth of the wetting layer, but these oscillations will damp rapidly upon

the change from two- to three-dimensional growth. This allows the determination of the thickness for the transition to islanding.

Many studies of the growth mode have relied on *ex situ* microscopy. Here, the morphology of the surface is examined using Nomarski interference contrast microscopy, SEM, or AFM. Films of various thicknesses can be examined in order to distinguish between the Stranski–Krastanov and Volmer–Weber modes.

Bean et al.[132] studied the growth mode of $Si_{1-x}Ge_x$ on Si(001) using Nomarski interference contrast microscopy. The objective of this study was to find the conditions under which SiGe could be grown on Si by MBE without islanding, to enable the growth of heterostructures and superlattices. They grew a series of $Si_{1-x}Ge_x$ layers, 100 nm thick, with various compositions and at various temperatures. Each layer was inspected using Nomarski interference contrast microscopy to determine the growth morphology. The map of Figure 7.90 was produced, in which it can be seen that planar layers may be grown over the entire compositional range at low temperatures.

Haffouz et al.[183] studied the growth mode for a GaN nucleation layer deposited on sapphire (0001) using SEM. Prior to growth, the sapphire wafers were subjected to a high-temperature Si/N treatment, involving exposure to SiH_4 and NH_3 at 1150°C, in the growth chamber. Then a 25–50 nm GaN nucleation layer was grown on *c*-plane sapphire by low-pressure MOVPE at 525°C. Following the growth of the nucleation layer, some samples were heated up to 1130°C–1150°C to simulate the temperature ramp that would be used prior to the growth of a thick GaN layer. Figure 7.91 shows SEM micrographs of the GaN samples. Figure 7.91a shows an as-grown nucleation layer, which has smooth surface morphology. The sample of Figure 7.91b was 25 nm and heated to 1130°C–1150°C for 60 s; the sample of Figure 7.91c was 50 nm thick and heated to 1130°C–1150°C for 60 s; the sample in Figure 7.91d was 50 nm thick and heated to 1130°C–1150°C for 30 s. It can be seen that the nucleation layer thickness and its temperature treatment control the size and density of the resulting GaN islands.

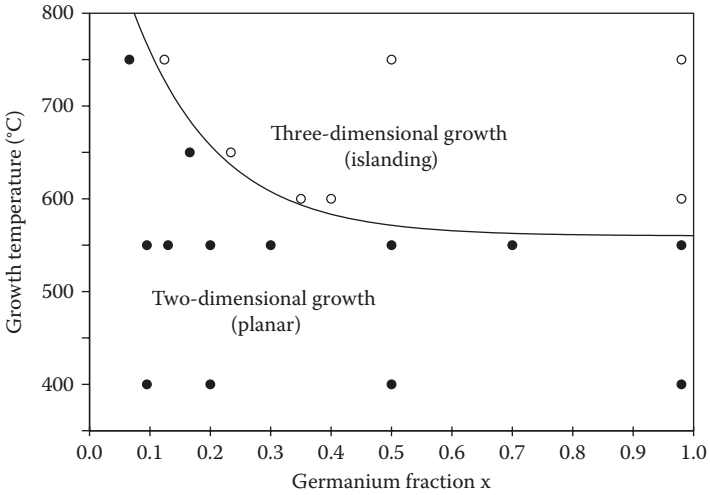


FIGURE 7.90
Film morphology for Ge_xSi_{1-x} on Si(001) grown with different growth temperatures and compositions, with a total thickness of 100 nm. The filled circles represent layers exhibiting two-dimensional growth, and the open circles are for samples that had rough morphology indicative of islanding. (Adapted from J. C. Bean et al., *J. Vac. Sci. Technol. A*, 2, 436 [1984]. With permission. Copyright 1984, American Vacuum Society.)

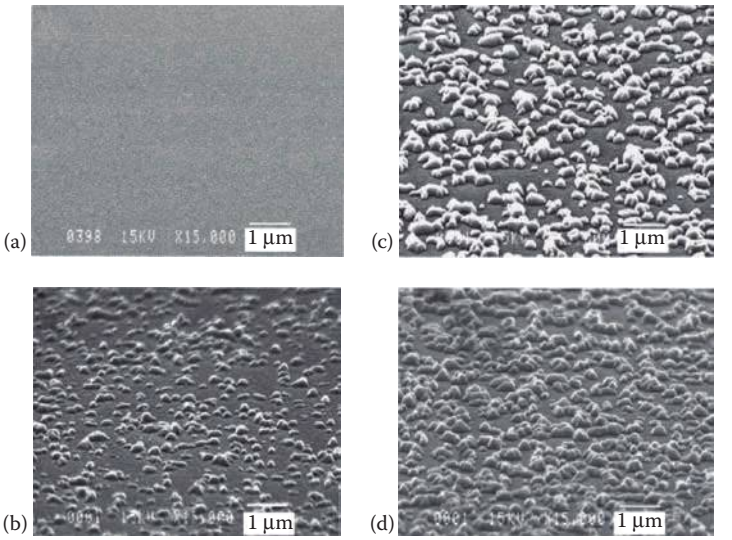


FIGURE 7.91
SEM micrographs of GaN nucleation layers grown on sapphire (0001) by low-pressure MOVPE following a Si/N treatment: (a) as-grown nucleation layer, (b) 25 nm nucleation layer that was heated to 1130°C–1150°C for 60 s, (c) 50 nm thick nucleation layer that was heated to 1130°C–1150°C for 60 s, and (d) 50 nm thick nucleation layer that was heated to 1130°C–1150°C for 30 s. (Reprinted from S. Haffouz et al., *Appl. Phys. Lett.*, 73, 1278 [1998]. With permission. Copyright 1998, American Institute of Physics.)

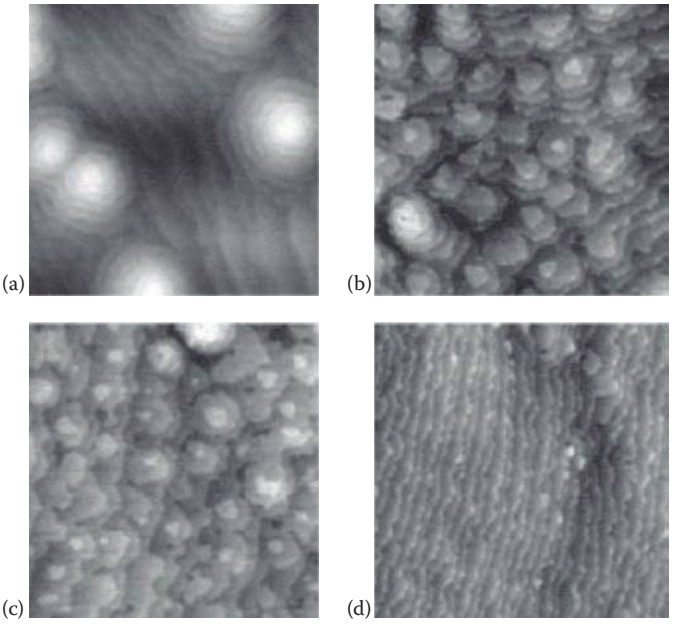


FIGURE 7.92
AFM micrographs (1 × 1 µm) showing the surfaces of 10 ML InGaN layers grown on GaN/sapphire (0001) by MOVPE at various NH_3 flow rates. All layers were grown at a temperature of 700°C–710°C, with a total pressure of 300 torr. The flows of trimethylgallium and trimethylindium were fixed at 2 and 100 sccm, respectively. The NH_3 flows were (a) 1 slm, (b) 3 slm, (c) 5 slm, and (d) 10 slm. (Reprinted from R. A. Oliver et al., *J. Appl. Phys.*, 97, 13707 [2005]. With permission. Copyright 2005, American Institute of Physics.)

AFM has proved to be invaluable in studies of quantum dots and quantum dot assembly processes, because of the atomic-scale resolution it affords. The reader is referred to Chapter 4, in which several examples of AFM micrographs clearly show the sizes, shapes, and distributions of nanometer-scale islands.

Oliver et al.[184] used AFM to study the growth modes of InGaN grown on GaN/sapphire (0001) by MOVPE. They studied the dependence of the growth modes on the temperature and source flows. Figure 7.92 shows some representative AFM micrographs obtained in this study, which show the capabilities of the method. The InGaN layers shown in Figure 7.92 were 10 ML thick, and grown at a temperature of 700°C–710°C, with a total pressure of 300 torr. The flows of trimethylgallium and trimethylindium were fixed at 2 and 100 sccm, respectively, while the NH_3 flow was varied. The growth morphology changes markedly as the NH_3 flow is varied from 1 to 10 slm. In Figure 7.92a, the surface shows large spiral mounds, separated by flat terraces that do not have monolayer islands. This is indicative of a modified step-flow growth mode. The samples of Figure 7.92b and c exhibit island growth with clear second-layer nucleation, indicating either a Volmer–Weber or Stranski–Krastanov growth mode. On the surface of the sample of Figure 7.92d, only terraces are observed, revealing a step-flow growth mode.

PROBLEMS

1. Suppose $Al_{0.2}Ga_{0.8}As$ is grown on GaAs(001). (a) Assuming the layer is pseudomorphic, determine the 004 x-ray Bragg angle difference, the 115 Bragg angle difference, and the change in the inclination of the (115) planes due to the tetragonal distortion. (b) Repeat the calculations with the assumption of a fully relaxed layer.
2. Calculate the 002, 004, and 006 Bragg angles for sapphire assuming Fe kα_1 radiation.
3. The 004 rocking curve is measured for a 1 μm thick $In_xGa_{1-x}As$ layer on an InP(001) substrate, at different azimuths. The peak separation varies from −400 to −600 arcsec. The FWHM for the epitaxial layer rocking curve varies from 180 to 220 arcsec. Estimate the composition of the layer, making reasonable assumptions about the state of strain in the $In_xGa_{1-x}As$.
4. ZnSe is grown heteroepitaxially on GaAs(001). For a 460 nm thick layer, the measured rocking curve width (FWHM) is 314 arcsec. (a) What is the expected rocking curve width for a perfect crystal of ZnSe with this thickness? Is the 460 nm layer relaxed, and if so, what is the approximate threading dislocation density? (b) Repeat for a 95 nm thick layer for which the measured 004 rocking curve width is 304 arcsec.
5. A layer of ZnS_xSe_{1-x} is grown on GaAs(001) and characterized by HRXRD. With the azimuth ψ set to zero, the projection of the x-ray incident beam is along the [110] direction in the surface of the sample. The (004) rocking curve peak separation was −90 arcsec at $\psi = 0°$ and also at $\psi = 180°$. The 044 peak separations were determined to be −150 arcsec at $\psi = 45°$and −120 arcsec at $\psi = 225°$, using the $\theta_B + \phi$ incidence for both measurements. Determine the 004 and 044 Bragg angles for the epitaxial layer, and its in-plane and out-of-plane lattice constants a and c. Assuming the elastic constants for the film are the same as for pure ZnSe, determine the relaxed lattice constant a_0. Using this value and Vegard's law, estimate the composition of the epitaxial layer.
6. The 004 rocking curve is measured for a vicinal Si(001) substrate at various azimuths. Determine the expected variation of the rocking curve width (FWHM)

assuming a perfect crystal with a 4° tilt and neglecting the instrumental broadening. Repeat assuming the instrument introduces a Gaussian broadening function with a width of 6 arcsec.

7. A 40-period 80 nm/100 nm $Al_{0.2}Ga_{0.8}As$/GaAs superlattice is grown on a GaAs(001) substrate. 004 rocking curves are measured using Cu kα radiation. If the entire structure is pseudomorphic, what is the position of the zero-order superlattice rocking curve peak, in relation to the substrate peak? What is the expected spacing for the superlattice peaks?
8. Assuming 100 keV electrons, what is the expected spacing between the 00 and 01 RHEED streaks for an unreconstructed GaAs surface if the sample-to-screen spacing is 30 cm?
9. An FTIR reflectance spectrum is measured for a 1 μm thick layer of $In_xGa_{1-x}P$ on GaAs(001). What is the expected fringe spacing (in wavenumber)?

References

1. W. L. Bragg, The diffraction of short electromagnetic waves by a crystal, *Proc. Cambridge Phil. Soc.*, 17, 43 (1913).
2. P. Kidd, XRD of gallium nitride and related compounds: Strain, composition and layer thickness [booklet], retrieved January 28, 2016, from http://www.panalytical.com/Technology-background/XRD-of-gallium-nitride-and-related-compounds-strain-composition-and-layer-thickness-booklet.htm.
3. N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Philadelphia: Holt, Rinehart and Winston, 1976).
4. *International Tables for X-Ray Crystallography*, ed. J. A. Ibers and W. C. Hamilton, Vol. IV (Birmingham, England: Kynoch Press, 1974).
5. S. N. Basu, Study of phase separation and ordering in InGaN alloys, *Electrochem Soc. Proc.*, 98-18, 15 (1999).
6. B. D. Cullity, *Elements of X-Ray Diffraction*, 2nd ed. (Reading, MA: Addison-Wesley, 1978), p. 139.
7. R. W. James, *The Crystalline State*, Vol. II: *The Optical Principles of the Diffraction of X-rays,* ed. Sir L. Bragg (London: G. Bell and Sons, 1948).
8. V. S. Speriosu, H. L. Glass, and T. Kobayashi, X-ray determination of strain and damage distributions in ion-implanted layers, *Appl. Phys. Lett.*, 34, 539 (1979); *Appl. Phys. Lett.*, 34, 539 (1979).
9. V. S. Speriosu, Kinematical x-ray diffraction in nonuniform crystalline films: Strain and damage distributions in ion-implanted garnets, *J. Appl. Phys.*, 52, 6094 (1981).
10. V. S. Speriosu, B. M. Paine, M. A. Nicolet, and H. L. Glass, X-ray rocking curve study of Si-implanted GaAs, Si, and Ge, *Appl. Phys. Lett.*, 40, 604 (1982).
11. V. S. Speriosu and C. H. Wilts, X-ray rocking curve and ferromagnetic resonance investigations of ion-implanted magnetic garnet, *J. Appl. Phys.*, 54, 3325 (1983).
12. V. S. Speriosu and T. Vreeland Jr., X-ray rocking curve analysis of superlattices, *J. Appl. Phys.*, 56, 3743 (1984).
13. V. S. Speriosu, M. A. Nicolet, J. L. Tandon, and Y. C. M. Yeh, Interfacial strain in $Al_xGa_{1-x}As$ layers on GaAs, *J. Appl. Phys.*, 57, 1377 (1985).
14. C. G. Darwin, The theory of x-ray reflection, *Phil Mag.*, 27, 315 (1914); C. G. Darwin, The theory of x-ray reflection, Part II, *Phil. Mag.*, 27, 675 (1914).
15. P. P. Ewald, Optics of crystals, *Ann. Physik*, 49, 1 (1916); *Ann. Physik*, 49, 117 (1916).

16. P. P. Ewald, Foundations of the optics of crystals. III. The crystal-optics of x-rays, *Ann. Physik*, 54, 519 (1916).
17. P. P. Ewald, Group velocity and phase velocity in x-ray crystal optics, *Acta Cryst.*, 11, 888 (1958).
18. M. von Laue, Dynamical theory of x-ray interference, *Ergeb. Exakt. Naturw.*, 10, 133 (1931).
19. B. W. Batterman and H. Cole, Dynamical diffraction of x rays by perfect crystals, *Rev. Mod. Phys.*, 36, 681 (1964).
20. W. H. Zachariasen, *Theory of X-Ray Diffraction in Crystals* (New York: John Wiley & Sons, 1950).
21. B. Klar and F. Rustichelli, Dynamical neutron diffraction by ideally curved crystals, *Il Nuovo Cimento*, 13B, 249 (1973).
22. S. Takagi, A dynamical theory of diffraction for a distorted crystal, *J. Phys. Soc. Jpn.*, 26, 1239 (1969).
23. S. Takagi, Dynamical theory of diffraction applicable to crystals with any kind of small distortion, *Acta Cryst.*, 15, 1311 (1962).
24. D. Taupin, Dynamical theory of x-ray reflection by deformed crystals, *C. R. Acad. Sci. (France)*, 256, 4881 (1963).
25. M. A. G. Halliwell, M. H. Lyons, and M. J. Hill, The interpretation of x-ray rocking curves from III-V semiconductor device structures, *J. Cryst. Growth*, 68, 523 (1984).
26. C. R. Wie, T. A. Tombrello, and T. Vreeland Jr., Dynamical x-ray diffraction from nonuniform crystalline films: Application to x-ray rocking curve analysis, *J. Appl. Phys.*, 59, 3743 (1986).
27. W. J. Bartels, Characterization of thin layers on perfect crystals with a high resolution x-ray diffractometer, *J. Vac. Sci. Technol. B*, 1, 338 (1983).
28. P. B. Rago, E. N. Suarez, F. C. Jain, and J. E. Ayers, Dynamical x-ray diffraction from ZnS_ySe_{1-y}/GaAs(001) multilayers and superlattices with dislocations, *J. Electron. Mater.*, 41, 2846 (2012).
29. P. B. Rago and J. E. Ayers, Dynamical X-ray diffraction from InGaAs heterostructures with dislocations, *J. Electron. Mater.*, 42, 2450 (2013).
30. P. B. Rago, F. C. Jain and J. E. Ayers, Effect of epilayer tilt on dynamical x-ray diffraction from uniform heterostructures with asymmetric dislocation densities, *J. Electron. Mater.*, 42, 3066 (2013).
31. P. B. Rago and J. E. Ayers, Mosaic crystal model for dynamical x-ray diffraction from step-graded In_xGa_{1-x}/As and In_xAl_{1-x}As/GaAs(001) metamorphic buffers and device structures, *Intl. J. High Speed Electron. Syst.*, in press (2016).
32. P. B. Rago and J. E. Ayers, Comparison of the phase-invariant and mosaic crystal models for dynamical x-ray diffraction from metamorphic In_xGa_{1-x}As/GaAs(001) structures, *J. Vac. Sci. Technol. B*, 33, 021204 (2015).
33. J. A. Prins, Die reflexion von rontgenstrahlen an absorbierenden idealen kristallen, *Z. Phys.*, 63, 477 (1930).
34. P. B. Hirsch and G. N. Ramachandran, Intensity of x-ray reflection from perfect and mosaic absorbing crystals, *Acta Cryst.*, 3, 187 (1950).
35. B. K. Tanner and M. J. Hill, Double axis x-ray diffractometry at glancing angles, *J. Phys. D*, 19, L229 (1986).
36. B. D. Cullity, *Elements of X-Ray Diffraction*, 2nd ed. (Reading, MA: Addison-Wesley, 1978), p. 511.
37. E. R. Cohen and B. N. Taylor, The 1986 adjustment of the fundamental physical constants, report of the Committee on Data for Science and Technology of the International Council of Scientific Unions (CODATA) Task Group on Fundamental Constants, CODATA Bulletin 63 (Elmsford, NY: Pergamon, 1986).
38. J. Ladell, W. Parish, and J. Taylor, Interpretation of diffractometer line profiles, *Acta Cryst.*, 12, 561 (1959).
39. H. Berger, Study of the Kα emission spectrum of copper, *X-Ray Spectrom.*, 15, 241 (1986).
40. J. Ayers and J. Ladell, Spectral widths of the Cu Kα lines, *Phys. Rev. A*, 37, 2404 (1988).
41. P. F. Fewster and N. L. Andrew, Determining the lattice relaxation in semiconductor layer systems by x-ray diffraction, *J. Appl. Phys.*, 74, 3121 (1993).
42. J. W. M. DuMond, Theory of the use of more than two successive x-ray crystal reflections to obtain increased resolving power, *Phys. Rev.*, 52, 872 (1937).

43. P. F. Fewster, A high-resolution multiple-crystal multiple-reflection diffractometer, *J. Appl. Cryst.*, 22, 64 (1989).
44. P. F. Fewster, Combining high-resolution x-ray diffractometry and topography, *J. Appl. Cryst.*, 24, 178 (1991).
45. T. Kujofsa and J. E. Ayers, Resolution of x-ray rocking curve measurements made with finite counting statistics, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 (2015).
46. J. Wang, H. Wu, R. So, Y. Liu, and M. H. Lie, Structure determination of indium-induced Si(111)-In-4 × 1 surface by LEED Patterson inversion, *Phys. Rev. B*, 72, 245324 (2005).
47. G. M. Nomarski, Microinterferometre differential a ondes polarisees, *J. Phys. Radium Paris*, 16, 9S (1955).
48. D. B. Williams and C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science* (Berlin: Springer, 2004).
49. K. R. Breen, P. N. Uppal, and J. S. Ahearn, Interface dislocation structures in $In_xGa_{1-x}As$/GaAs mismatched epitaxy, *J. Vac. Sci. Technol. B*, 7, 758 (1989).
50. G. Binnig, C. F. Quate, and Ch. Gerber, Atomic force microscope, *Phys. Rev. Lett.*, 56, 930 (1986).
51. A. Jasik, K. Kosiel, W. Strupinski, and M. Wesolowski, Influence of covering on critical thickness of strained $In_xGa_{1-x}As$ layer, *Thin Solid Films*, 412, 50 (2002).
52. M. Tachikawa and H. Mori, Dislocation generation of GaAs on Si in the cooling stage, *Appl. Phys. Lett.*, 56, 2225 (1990).
53. T. Ohno, A. Ohki, and T. Matsuoka, Surface cleaning with hydrogen plasma for low-defect-density ZnSe homoepitaxial growth, *J. Vac. Sci. Technol. A*, 16, 2539 (1998).
54. J. C. Campbell, S. M. Abbott, and A. G. Dentai, A comparison of 'normal' lasers and lasers exhibiting light jumps, *J. Appl. Phys.*, 51, 4010 (1980).
55. H. Kasano, Regular compositional steps generated in $GaAs_{1-x}P_x$ VPE layers, *J. Appl. Phys.*, 51, 4178 (1980).
56. Y. Huang, P. Y. Yu, H. Lee, and S. Wang, Characterization of GaAs film grown on Si substrate by photoluminescence at 77K, *Appl. Phys. Lett.*, 52, 579 (1988).
57. P. L. Gourley, R. M. Biefeld, and L. R. Dawson, Elimination of dark line defects in lattice-mismatched epilayers through use of strained-layer superlattices, *Appl. Phys. Lett.*, 47, 482 (1985).
58. K. Matsuda, T. Saiki, H. Saito, and K. Nishi, Room-temperature photoluminescence of self-assembled $In_{0.5}Ga_{0.5}As$ single quantum dots by using highly sensitive near-field scanning optical microscope, *Appl. Phys. Lett.*, 76, 73 (2000).
59. T. Saiki, S. Mononobe, M. Ohtsu, N. Saito, and J. Kusano, Tailoring a high-transmission fiber probe for photon scanning tunneling microscope, *Appl. Phys. Lett.*, 68, 2612 (1996).
60. P. L. Gourley, T. J. Drummond, and B. L. Doyle, Dislocation filtering in semiconductor superlattices with lattice-matched and lattice-mismatched layer materials, *Appl. Phys. Lett.*, 49, 1101 (1986).
61. Y. Ohizumi, T. Tsuruoka, and S. Ushioda, Formation of misfit dislocations in GaAs/InGaAs multiquantum wells observed by photoluminescence microscopy, *J. Appl. Phys.*, 92, 2385 (2002).
62. W. Braun, L. Daweritz, K. H. Ploog, New model for reflection high-energy electron diffraction intensity oscillations, *J. Vac. Sci. Technol. B*, 16, 2404 (1998).
63. Y. Yu. Davydov, N. S. Averkiev, I. N. Goncharuk, D. K. Nelson, I. P. Nikitina, A. S. Polkovnikov, A. N. Smirnov, and M. A. Jacobson, Raman and photoluminescence studies of biaxial strain in GaN epitaxial layers grown on 6H-SiC, *J. Appl. Phys.*, 82, 5097 (1997).
64. E. Estop, A. Izrael, and M. Sauvage, Double-crystal spectrometer measurements of lattice parameters and x-ray topography on heterojunctions GaAs-$Al_xGa_{1-x}As$, *Acta Cryst. A*, 32, 627 (1976).
65. J. Hornstra and W. J. Bartels, Determination of the lattice constant of epitaxial layers of III-V compounds, *J. Cryst. Growth*, 44, 513 (1978).
66. W. J. Bartels and W. Nijman, X-ray double-crystal diffractometry of $Ga_{1-x}Al_xAs$ epitaxial layers, *J. Cryst. Growth*, 44, 518 (1978).
67. A. T. Macrander, G. P. Schwartz, and G. J. Gualtieri, X-ray and Raman characterization of AlSb/GaSb strained layer superlattices and quasiperiodic Fibonacci lattices, *J. Appl. Phys.*, 64, 6733 (1988).

68. A. Leiberich and J. Levkoff, A double crystal x-ray diffraction characterization of $Al_xGa_{1-x}As$ grown on an offcut GaAs(100) substrate, *J. Vac. Sci. Technol. B*, 8, 422 (1990).
69. A. Krost, G. Bauer, and J. Woitok, in *Optical Characterization of Epitaxial Semiconductor Layers*, ed. G. Bauer and W. Richter (Berlin: Springer, 1996), p. 287.
70. B. E. Warren, *X-Ray Diffraction* (Reading, MA: Addison-Wesley, 1969), p. 353.
71. E. Koppensteiner, P. Hamberger, G. Bauer, A. Pesek, H. Kibbel, H. Presting, and E. Kasper, X-ray diffraction investigation of single step and step-graded SiGe alloy buffers for the growth of short-period Si_mGe_n superlattices using reciprocal space mapping, *Appl. Phys. Lett.*, 62, 1783 (1993).
72. D. Lee, M. S. Park, Z. Tang, H. Luo, R. Beresford, and C. R. Wie, Characterization of metamorphic $In_xAl_{1-x}As$/GaAs buffer layers using reciprocal space mapping, *J. Appl. Phys.*, 101, 063523 (2007).
73. M. Chauveau, Y. Androussi, A. Lefebvre, J. Di Persio, and Y. Cordier, Indium content measurements in metamorphic high electron mobility transistor structures by combination of x-ray reciprocal space mapping and transmission electron microscopy, *J. Appl. Phys.*, 93, 4219 (2003).
74. G. Belenky, D. Wang, Y. Lin, D. Donetsky, G. Kipshidze, L. Shterengas, D. Westerfeld, W. L. Sarney, and S. P. Svensson, Metamorphic InAsSb/AlInAsSb heterostructures for optoelectronic applications, *Appl. Phys. Lett.*, 102, 111108 (2013).
75. S. Pereira, M. R. Correia, E. Pereira, K. P. O'Donnell, E. Alves, A. D. Segueira, N. Franco, I. M. Watson, and C. J. Deatcher, Strain and composition distributions in wurtzite InGaN/GaN layers extracted from x-ray reciprocal space mapping, *Appl. Phys. Lett.*, 80, 3913 (2002).
76. S. Pereira, M. R. Coreira, E. Pereira, C. Trager-Cowan, F. Sweeney, K. P. O'Donnell, E. Alves, N. Franco, and A. D. Sequeira, Structural and optical properties of InGaN/GaN layers close to the critical layer thickness, *Appl. Phys. Lett.*, 81, 1207 (2002).
77. M. A. Moram, C. F. Johnston, J. L. Hollander, M. J. Kapers, and C. J. Humphreys, Understanding x-ray diffraction of nonpolar gallium nitride films, *J. Appl. Phys.*, 105, 113501 (2009).
78. G. G. Stoney, The tension of metallix films deposited by electrolysis, *Proc. R. Soc. A*, 82, 172 (1909).
79. L. B. Freund, Some elementary connections between curvature and mismatch strain in compositionally graded thin films, *J. Mech. Phys. Solids*, 44, 723 (1996).
80. L. B. Freund, Substrate curvature due to thin film mismatch strain in the nonlinear deformation range, *J. Mech. Phys. Solids*, 48, 1159 (2000).
81. L. B. Freund, J. A. Floro, and E. Chason, Extensions of the Stoney formula for substrate curvature to configurations with thin substrates or large deformations, *Appl. Phys. Lett.*, 74, 1987 (1999).
82. J. A. Floro, E. Chason, S. R. Lee, R. D. Twesten, R. Q. Hwang, and L. B. Freund, Real-time stress evolution during SiGe heteroepitaxy—Dislocations, islanding, and segregation, *J. Electron. Mater.*, 26, 969 (1997).
83. J. A. Floro, E. Chason, R. D. Twesten, R. Q. Hwang, and L. B. Freund, Kinetics of stress relaxation due to coherent island formation during $Si_{0.8}Ge_{0.2}$/Si epitaxial growth, *Phys. Rev. Lett.*, 79, 3946 (1997).
84. J. A. Floro, G. A. Lucadamo, E. Chason, L. B. Freund, M. Sinclair, R. D. Twesten, and R. Q. Hwang, SiGe island shape transitions induced by elastic repulsion, *Phys. Rev. Lett.*, 80, 4717 (1998).
85. B. Pérez Rodríguez and J. Mirecki Millunchick, The role of morphology in the relaxation of strain in InGaAs/GaAs, *J. Cryst. Growth*, 264, 64 (2004).
86. R. Beresford, C. Lynch, and E. Chason, Kinetics of dislocation-mediated strain relaxation in InGaAs/GaAs heteroepitaxy, *J. Cryst. Growth*, 251, 106 (2003).
87. C. Lynch, E. Chason, R. Beresford, and S. K. Hong, Influence of growth flux and surface supersaturation on InGaAs/GaAs strain relaxation, *Appl. Phys. Lett.*, 84, 1085 (2004).
88. C. Lynch, E. Chason and R. Beresford, Enhanced strain relaxation rate of InGaAs by adatom-assisted dislocation kink nucleation, *J. Vac. Sci. Technol. B*, 23, 1166 (2005).

89. R. Beresford, K. Tetz, J. Yin, E. Chason and M. U. González, Metastability of InGaAs/GaAs probed by *in situ* optical stress sensor, *J. Vac. Sci. Technol. B*, 19, 1572 (2001).
90. J. Petruzzello, B. L. Greenberg, D. A. Cammack, and R. Dalby, Structural properties of the ZnSe/GaAs system grown by molecular beam epitaxy, *J. Appl. Phys.*, 63, 2299 (1988).
91. A. Mazuelas, L. Gonzalez, F. A. Ponce, L. Tapfer, and F. Briones, Critical thickness determination of InAs, InP and GaP on GaAs by x-ray interference effect and transmission electron microscopy, *J. Cryst. Growth*, 131, 465 (1993).
92. H. Yamaguchi, J. G. Belk, X. M. Zhang, J. L. Sudijono, M. R. Fahy, T. S. Jones, D. W. Pashley, and B. A. Joyce, Atomic-scale imaging of strain relaxation via misfit dislocations in highly mismatched semiconductor heteroepitaxy: InAs/GaAs(111)A, *Phys. Rev. B*, 55, 1337 (1997).
93. T. Yao, The effect of lattice misfit on lattice parameters and photoluminescence properties of atomic layer epitaxy grown ZnSe on (100) GaAs substrates, *Jpn. J. Appl. Phys.*, 25, L544 (1986).
94. M. J. Reed, N. A. El-Masry, C. A. Parker, J. C. Roberts, and S. M. Bedair, Critical layer thickness determination of GaN/InGaN/GaN double heterostructures, *Appl. Phys. Lett.*, 77, 4121 (2000).
95. P. L. Gourley, I. J. Fritz, and L. R. Dawson, Controversy of critical layer thickness for InGaAs/GaAs strained-layer epitaxy, *Appl. Phys. Lett.*, 52, 377 (1988).
96. I. J. Fritz, P. L. Gourley, and L. R. Dawson, Dependence of critical layer thickness on strain for $In_xGa_{1-x}As$/GaAs strained-layer superlattices, *Appl. Phys. Lett.*, 46, 967 (1985).
97. I. J. Fritz, P. L. Gourley, and L. R. Dawson, Critical layer thickness in $In_{0.2}Ga_{0.8}As$/GaAs single strained quantum well structures, *Appl. Phys. Lett.*, 51, 1004 (1987).
98. T. Reisinger, M. J. Kastner, K. Wolf, E. Steinkirchner, W. Hackl, H. Stanzl, and W. Gebhardt, Critical thickness determination of II-VI semiconductors by RHEED and x-ray diffraction, *Mater. Sci. Forum*, 182–184, 147 (1995).
99. P. J. Orders and B. F. Usher, Determination of critical layer thickness in $In_xGa_{1-x}As$/GaAs heterostructures by x-ray diffraction, *Appl. Phys. Lett.*, 50, 980 (1987).
100. Y. C. Chen and P. K. Bhattacharya, Determination of critical layer thickness and strain tensor in $In_xGa_{1-x}As$/GaAs quantum-well structures by x-ray diffraction, *J. Appl. Phys.*, 73, 7389 (1993).
101. Y. Kohama, Y. Fukuda, and M. Seki, Determination of the critical layer thickness of $Si_{1-x}Ge_x$/Si heterostructures by direct observation of misfit dislocations, *Appl. Phys. Lett.*, 52, 380 (1988).
102. H.-J. Gossman, G. P. Schwartz, B. A. Davidson, and G. J. Gultieri, Strain and critical thickness in GaSb(001)/AlSb, *J. Vac. Sci. Technol. B*, 7, 764 (1989).
103. R. People and J. C. Bean, Calculation of critical layer thickness versus lattice mismatch for Ge_xSi_{1-x}/Si strained-layer heterostructures, *Appl. Phys. Lett.*, 47, 322 (1985); *Appl. Phys. Lett.*, 49, 229 (1986).
104. A. Fischer, H. Kuhne, and H. Richter, New approach in equilibrium theory for strained layer relaxation, *Phys. Rev. Lett.*, 73, 2712 (1994).
105. I. J. Fritz, Role of experimental resolution in measurements of critical layer thickness for strained-layer epitaxy, *Appl. Phys. Lett.*, 51, 1080 (1987).
106. J. E. Ayers, S. K. Ghandhi, and L. J. Schowalter, Crystallographic tilting of heteroepitaxial layers, *J. Cryst. Growth*, 113, 430 (1991).
107. S. K. Ghandhi and J. E. Ayers, Strain and misorientation in GaAs grown on Si(001) by organometallic epitaxy, *Appl. Phys. Lett.*, 53, 1204 (1988).
108. X. G. Zhang, D. W. Parent, P. Li, A. Rodriguez, G. Zhao, J. E. Ayers, and F. C. Jain, X-ray rocking curve analysis of tetragonally distorted ternary semiconductors on mismatched (001) substrates, *J. Vac. Sci. Technol. B*, 18, 1375 (2000).
109. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers. I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).
110. W. D. Laidig, C. K. Peng, and Y. F. Lin, Effects of strain and layer thickness on the growth of $In_xGa_{1-x}As$-GaAs strained-layer superlattices, *J. Vac. Sci. Technol. B*, 2, 181 (1984).
111. T. Kujofsa, S. Cheruku, W. Yu, B. Outlaw, S. Xhurxhi, D. Sidoti, F. Obst, et al., Critical layer thickness in II-VI/GaAs heterostructures and the role of finite experimental resolution, presented at U.S. Workshop on the Physics and Chemistry of II-VI Materials, Chicago (2011).
112. K. Kamigaki, H. Sakashita, H. Kato, M. Nakayama, N. Sano, and H. Terauchi, X-ray study of misfit strain relaxation in lattice-mismatched heterojunctions, *Appl. Phys. Lett.*, 49, 1071 (1986).

113. X. G. Zhang, P. Li, D. W. Parent, G. Zhao, J. E. Ayers, and F. C. Jain, Comparison of x-ray diffraction methods for determination of the critical layer thickness for dislocation multiplication, *J. Electron. Mater.*, 28, 553 (1999).
114. P. Scherrer, Bestimmung der Grosse und der inneren struktur von kolloidteilchen mittels rontgenstrahlen, *Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen, Mathematisch-Physikalische Klasse*, 1918, 98 (1918).
115. B. D. Cullity, *Elements of X-Ray Diffraction* (New York: Addison Wesley, 1978), p. 102.
116. B. Post, J. Nicolosi, and J. Ladell, Experimental procedures for the determination of the invariant phases of centrosymmetric crystals, *Acta Cryst. A*, 40, 684 (1984).
117. C. Kim, I. K. Robinson, J. Myoung, K. Shim, M.-C. Yoo, and K. Kim, Critical thickness of GaN thin films on sapphire (0001), *Appl. Phys. Lett.*, 69, 2358 (1996).
118. W. J. Bartels, J. Hornstra, and D. J. W. Lobeek, X-ray diffraction of multilayers and superlattices, *Acta. Cryst. A*, 42, 539 (1986).
119. L. Tapfer and K. Ploog, X-ray interference in ultrathin epitaxial layers: A versatile method for the structural analysis of single quantum wells and heterointerfaces, *Phys. Rev.* B, 40, 9802 (1989).
120. P. Gay, P. B. Hirsch, and A. Kelly, The estimation of dislocation densities in metals from x-ray data, *Acta. Met.*, 1, 315 (1953).
121. M. J. Hordon and B. L. Averbach, X-ray measurements of dislocation density in deformed copper and aluminum single crystals, *Acta Met.*, 9, 237 (1961).
122. J. E. Ayers, The measurement of threading dislocation densities in semiconductor crystals by x-ray diffraction, *J. Cryst. Growth*, 135, 71 (1994).
123. A. Kamata and H. Mitsuhashi, Characterization of ZnSSe on GaAs by etching and x-ray diffraction, *J. Cryst. Growth*, 142, 31 (1994).
124. K. Kohra, H. Hashizume, and J.-I. Yoshimura, X-ray diffraction topography utilizing double-crystal arrangement of (+,+) or non-parallel (+,–) setting, *Jpn. J. Appl. Phys.*, 9, 1029 (1970).
125. C. B. O'Donnell, G. Lacey, G. Horsburgh, A. G. Cullis, C. R. Whitehouse, P. J. Parbrook, W. Meredith, I. Galbraith, P. Möck, K. A. Prior, and B. C. Cavenett, Measurements by x-ray topography of the critical thickness of ZnSe grown on GaAs, *J. Cryst. Growth*, 184/185, 95 (1998).
126. S. P. Edirisinghe, A. E. Staton-Bevan, and R. Grey, Relaxation mechanisms in single $In_xGa_{1-x}As$ epilayers grown on misoriented GaAs $(\bar{1}\bar{1}\bar{1})$ B substrates, *J. Appl. Phys.*, 82, 4870 (1997).
127. C. A. Parker, J. C. Roberts, S. M. Bedair, M. J. Reed, S. X. Liu, and N. A. El-Masry, Determination of the critical layer thickness in the InGaN/GaN heterostructures, *Appl. Phys. Lett.*, 75, 2776 (1999).
128. B. Elman, E. S. Koteles, P. Melman, C. Jagannath, J. Lee, and D. Dugger, *In situ* measurements of critical layer thickness and optical studies of InGaAs quantum wells grown on GaAs substrates, *Appl. Phys. Lett.*, 55, 1659 (1989).
129. C. W. Snyder, B. G. Orr, D. Kessler, and L. M. Sander, Effect of strain on surface morphology in highly strained InGaAs films, *Phys. Rev. Lett.*, 66, 3032 (1991).
130. S. Ohkouchi, N. Ikoma, and M. Tamura, Growth mode transition processes in a GaAs/InP system studied by scanning tunneling microscopy, *J. Electron. Mater.*, 25, 439 (1996).
131. D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L. E. Swartz, Early stages of growth of GaAs on Si observed by scanning tunneling microscopy, *Appl. Phys. Lett.*, 57, 2419 (1990).
132. J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, Ge_xSi_{1-x}/Si strained-layer superlattice grown by molecular beam epitaxy, *J. Vac. Sci. Technol. A*, 2, 436 (1984).
133. B. Greenberg and J. Ladell, Modulation of Renninger scan intensity: A new x-ray technique to characterize epitaxial structures, *Appl. Phys. Lett.*, 50, 436 (1987).
134. M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, Dislocation morphology in graded heterojunctions: $GaAs_{1-x}P_x$, *J. Mater. Sci.*, 4, 223 (1969).
135. F. L. Vogel, W. G. Pfann, II, E. Corey, and E. E. Thomas, Observation of dislocations in lineage boundaries in germanium, *Phys. Rev.*, 90, 489 (1953).
136. S. Ha, W. M. Vetter, M. Dudley, and M. Skowronski, A simple mapping method for elementary screw dislocations in homoepitaxial SiC layers, *Mater. Sci. Forum*, 389–393, 443 (2002).

137. Z. Zhang, Y. Gao, and T. S. Sudarshan, Delineating structural defects in highly doped n-type 4H-SiC substrates using a combination of thermal diffusion and molten KOH etching, *Electrochem. Solid State Lett.*, 7, G264 (2004).
138. K. Ishida, *Proc. Mater. Res. Soc. Symp.*, 19, 133 (1987).
139. K. Ishida, M. Akiyama, and S. Nishi, Misfit and threading dislocations in GaAs layers grown on Si substrates by MOCVD, *Jpn. J. Appl. Phys.*, 26, L163 (1987).
140. D. J. Stirland, Quantitative defect etching of GaAs on Si: Is it possible? *Appl. Phys. Lett.*, 53, 2432 (1988).
141. W. C. Dash, Copper precipitation on dislocations in silicon, *J. Appl. Phys.*, 27, 1193 (1956).
142. E. Sirtl and A. Adler, Chromsäure-Flussäure als spezifisches system zur Ätzgruben entwicklung auf silizium, *Z. Met.*, 52, 529 (1961).
143. F. Secco d'Aragona, Dislocation etch for (100) planes in silicon, *J. Electrochem. Soc.*, 119, 948 (1972).
144. M. W. Jenkins, A new preferential etch for defects in silicon crystals, *J. Electrochem. Soc.*, 124, 757 (1977).
145. J. G. Grabmaier and C. B. Watson, Dislocation etch pits in single crystal GaAs, *Phys. Status Solidi*, 32, K13 (1969).
146. J. Angilello, R. M. Potemski, and G. R. Woolhouse, Etch pits and dislocations in {100} GaAs wafer, *J. Appl. Phys.*, 46, 2315 (1975).
147. K. Akita, T. Kusunoki, S. Komiya, and T. Kotani, Observation of etch pits in InP by new etchants, *J. Cryst. Growth*, 46, 783 (1979).
148. F. Kuhn-Kuhnenfeld, D. Kirsten, and M. Maier, Etching agent for III/V semiconductors, U.S. Patent US 4100014 A (1978).
149. H. C. Gatos and M. C. Lavine, Dislocation etch pits on the {111} and $\{\bar{1}\bar{1}\bar{1}\}$ surfaces of InSb, *J. Appl. Phys.*, 31, 743 (1960).
150. Y. G. Chai and R. Chow, Molecular beam epitaxial growth of lattice-mismatched $In_{0.77}Ga_{0.23}As$ on InP, *J. Appl. Phys.*, 53, 1229 (1982).
151. T. J. de Lyon, D. Rajavel, S. M. Johnson, and C. A. Cockrum, Molecular-beam epitaxial growth of CdTe(112) on Si(112) substrates, *Appl. Phys. Lett.*, 66, 2119 (1995).
152. W. J. Everson, C. K. Ard, J. L. Sepich, B. E. Dean, and G. T. Neugebauer, Etch pit characterization of CdTe and CdZnTe substrate for use in mercury cadmium telluride epitaxy, *J. Electron. Mater.*, 24, 505 (1995).
153. J.-S. Chen, Etchant for revealing dislocations in II-VI compounds, U.S. Patent 4,897,152 (1990).
154. T. Koyama, T. Yodo, H. Oka, K. Yamashita, and T. Yamasaki, Growth of ZnSe single crystals by iodine transport, *J. Cryst. Growth*, 91, 639 (1988).
155. G. D. U'Ren, M. S. Goorsky, G. Meis-Haugen, K. K. Law, T. J. Miller, and K. W. Haberern, Defect characterization of etch pits in ZnSe based epitaxial layers, *Appl. Phys. Lett.*, 69, 1089 (1996).
156. C. C. Chu, T. B. Ng, J. Han, G. C. Hua, R. L. Gunshor, E. Ho, E. L. Warlick, and L. A. Kolodziejski, Reduction of structural defects in II-VI blue green laser diodes, *Appl. Phys. Lett.*, 69, 1089 (1996).
157. L. Zeng, B. X. Yang, B. Shewareged, M. C. Tamargo, J. Z. Wan, F. H. Pollak, E. Snoeks, and L. Zhao, Determination of defect density in ZnCdMgSe layers grown on InP using a chemical etch, *J. Appl. Phys.*, 82, 3306 (1997).
158. Y. Ono, Y. Iyechika, T. Takada, K. Inui, and T. Matsye, Reduction of etch pit density on GaN by InGaN-strained SQW, *J. Cryst. Growth*, 189, 133 (1998).
159. C.-C. Tsai, C.-S. Chang, and T.-Y. Chen, Low-etch-pit-density GaN substrates by regrowth on free-standing GaN films, Appl. Phys. Lett., 80, 3718 (2002).
160. T. Kozawa, T. Kachi, T. Ohwaki, Y. Taga, N. Koide, and M. Koike, Dislocation etch pits in GaN epitaxial layers grown on sapphire substrates, *J. Electrochem. Soc.*, 143, L17 (1996).
161. S. J. Rosner, E. C. Carr, M. J. Ludowise, G. Girolami, and H. I. Erikson, Correlation of cathodoluminescence inhomogeneity with microstructural defects in epitaxial GaN grown by metal-organic chemical-vapor deposition, *Appl. Phys. Lett.*, 70, 420 (1997).
162. J. E. Ayers, L. J. Schowalter, and S. K. Ghandhi, Post-growth thermal annealing of GaAs on Si(001) grown by organometallic vapor phase epitaxy, *J. Cryst. Growth*, 125, 329 (1992).

163. V. M. Kaganer, O. Brandt, A. Trampert, and K. H. Ploog, X-ray diffraction peak profiles from threading dislocations in GaN epitaxial films, *Phys. Rev.* B, 72, 45423 (2005).
164. C. O. Dunn and E. F. Koch, Comparison of dislocation densities of primary and secondary recrystallization grains of Si-Fe, *Acta Met.*, 5, 548 (1957).
165. B. E. Warren, Theory presented in the course on x-ray and crystal physics, Massachusetts Institute of Technology (1957).
166. W. F. Flanagan, Effect of substructure on the cleavage fracture of iron crystals, ScD dissertation, Massachusetts Institute of Technology (1959).
167. O. Yastrubchak, T. Wosinski, J. Z. Domagala, E. Lusakowska, T. Figielski, B. Pecz, and A. L. Toth, Misfit strain anisotropy in partially relaxed lattice-mismatched InGaAs/GaAs heterostructures, *J. Phys. Cond. Matter*, 16, S1 (2004).
168. B. Yarlagadda, A. Rodriguez, P. Li, R. Velampati, J. F. Ocampo, E. N. Suarez, P. B. Rago, D. Shah, J. E. Ayers, and F. C. Jain, X-ray characterization of dislocation density asymmetries in heteroepitaxial semiconductors, *Appl. Phys. Lett.*, 92, 202103 (2008).
169. M. R. Gokhale, K. X. Bao, P. D. Healey, F. C. Jain, and J. E. Ayers, Factors influencing low-temperature photo-assisted OMVPE growth of ZnSe, *J. Cryst. Growth*, 165, 25 (1996).
170. X. G. Zhang, S. Kalisetty, J. Robinson, G. Zhao, D. W. Parent, J. E. Ayers, and F. C. Jain, Structural properties of ZnS_ySe_{1-y}/ZnSe/GaAs(001) heterostructures grown by photoassisted metalorganic vapor phase epitaxy, *J. Electron. Mater.*, 26, 697 (1997).
171. E. Koppensteiner, A. Schuh, G. Bauer, V. Holy, G. P. Watson, and E. A. Fitzgerald, Determination of threading dislocation density in hetero-epitaxial layers by diffuse x-ray scattering, *J. Phys. D*, 28, A114 (1995).
172. C. J. K. Richardson, L. He, and S. Kanakaraju, Metamorphic growth of II-V semiconductor bicrystals, *J. Vac. Sci. Technol. B*, 29, 03C126-1 (2011).
173. V. M. Kaganer, R. Kohler, M. Schmidbauer, R. Opitz, and B. Jenichen, X-ray diffraction peaks due to misfit dislocations in heteroepitaxial structures, *Phys. Rev.* B, 55, 1793 (1997).
174. B. K. Tanner and M. J. Hill, X-ray double crystal diffractometry of multiple and very thin heteroepitaxial layers, *Adv. X-Ray Anal.*, 29, 337 (1986).
175. P. F. Fewster and C. J. Curling, Composition and lattice-mismatch measurement of thin semiconductors by x-ray diffraction, *J. Appl. Phys.*, 62, 4154 (1987).
176. I. Kim, S.-W. Ryu, B.-D. Choe, H.-D. Kim, and W. G. Jeong, Matrix method for the x-ray rocking curve simulation, *J. Appl. Phys.*, 83, 3932 (1998).
177. P. F. Fewster, N. L. Andrew, O. H. Hughes, C. Staddon, C. T. Foxon, A. Bell, T. S. Cheng, T. Wang, S. Sakai, K. Jacobs, and I. Moerman, X-ray studies of group III-nitride quantum wells with high quality interfaces, *J. Vac. Sci. Technol. B*, 18, 2300 (2000).
178. C. R. Wie and H. M. Kim, Kinematical x-ray diffraction model with a new boundary condition for analysis of Bragg-peak profiles of layered crystals, *J. Appl. Phys.*, 69, 6406 (1991).
179. M. A. Krivoglaz and K. P. Ryaboshapka, Theory of x-ray scattering by crystals containing dislocation walls, *Fiz. Metal. Metalloved.*, 15, 18 (1963).
180. L. E. Levine and R. Thomson, X-ray scattering by dislocations in crystals. General theory and application to screw dislocations, *Acta Cryst.*, 53, 590 (1997).
181. M. S. Abrahams, J. Blanc, and C. J. Buiocchi, Like-sign asymmetric dislocations in zinc-blende structure, *Appl. Phys. Lett.*, 21, 185 (1972).
182. Z. Jiang, W. Wang, H. Gao, L. Liu, H. Chen and J. Zhou, Strain relaxation and surface morphology of high indium content InAlAs metamorphic buffers with reverse step, *Appl. Surf. Sci.*, 254, 5241 (2008).
183. S. Haffouz, H. Larèche, P. Vennéguès, P. de Mierry, B. Beaumont, F. Omnès, and P. Gibart, The effect of the Si/N treatment of a nitrided sapphire surface on the growth mode of GaN in low-pressure metalorganic vapor phase epitaxy, *Appl. Phys. Lett.*, 73, 1278 (1998).
184. R. A. Oliver, M. J. Kappers, C. J. Humphreys, and G. A. D. Biggs, Growth modes in heteroepitaxy of InGaN on GaN, *J. Appl. Phys.*, 97, 13707 (2005).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

8

Defect Engineering in Heteroepitaxial Material

8.1 Introduction

Defect engineering in heteroepitaxial layers refers to efforts to control the densities, types, or arrangements of defects, especially dislocations. The common approaches to defect engineering involve the use of buffer layers, dislocation compensation, patterned substrates, patterning and annealing, epitaxial lateral overgrowth (ELO), or compliant substrates. Some of these techniques, such as buffer layers, patterning with annealing, and ELO, are intended to remove existing defects from relaxed heteroepitaxial layers. Others, such as reduced area growth, nanoheteroepitaxy (NHE), and compliant substrates, are designed to prevent the introduction of dislocations in the first place. This chapter presents the theory and practice of these common defect engineering approaches.

8.2 Buffer Layer Approaches and Virtual Substrates

It has already been shown that buffer layers may be used to engineer several properties of a top device layer, including the surface dislocation density. These approaches have been discussed at length in Chapter 6 and the reader is referred to that material for treatments of uniform, linear, step-graded, and nonlinearly graded buffers and their threading dislocation behavior.

A particular application of this approach that is worth mentioning here is the virtual substrate (VS). If the buffer layer is designed to be lattice matched to the device layer, then this device layer may be grown on top of it without the introduction of new dislocations. In principle, then, the use of a sufficiently thick buffer layer will allow the growth of a device layer with a low dislocation density on a convenient lattice-mismatched substrate.

An example of this is a thick epitaxial layer of GaN on a sapphire substrate, which can serve as a virtual GaN substrate, even though real GaN substrates are difficult to produce with high crystal quality. If the GaN buffer layer is very thick, it will behave as a real GaN substrate in some, but not all, respects. It is expected that the VS will be relaxed at the growth temperature, and that it will have a low dislocation density. But, unless the thick buffer is exfoliated from its substrate, it will be constrained to mimic the thermal expansion of the substrate. All the same, VSs open up possibilities for metamorphic materials, such as ternary or quaternary semiconductors, which cannot be readily manufactured in bulk.

8.2.1 Dislocation Compensation

In compositionally graded metamorphic buffer layers, it has been shown that the introduction of a mismatched interface can bend over and remove threading dislocations. This specialized area of application was not discussed in Chapter 6 and is therefore considered here. The dislocation compensation mechanism involves the bending over of threading dislocations associated with misfit segments of one sense by those associated with the opposite sense, thus allowing removal of threading dislocations from the surface. Here, the sense refers to the direction of the Burgers vector. We can arbitrarily assign positive and negative sense to dislocations that relax compressive and tensile strain, respectively. Qualitatively, a compressive (tensile) interface can be used for dislocation compensation in a structure with a tensile (compressive) buffer. Partial dislocation compensation by the insertion of a mismatched interface has been verified experimentally in multilayered ZnS_ySe_{1-y}/GaAs(001),[1] $Si_{1-x}Ge_x$/Si(001),[2] and $In_xGa_{1-x}As$/GaAs(001)[3,4] heterostructures.

For example, dislocation compensation may be realized at an overshoot interface between a graded buffer layer and a device layer. In this situation, for a given change in composition (overshoot) at the buffer–device layer interface, there is an optimum buffer layer thickness that minimizes the dislocation density, and also, for a given buffer layer thickness, there is an optimum overshoot. In practical structures, perfect compensation may be unachievable because some dislocations could be sessile or pinned; however, in principle it may be possible to remove all mobile threading dislocations by adjusting the overshoot or buffer layer thickness. In practice, this goal may be considered to have been met if the ending threading dislocation is comparable to that of the substrate (10 to $10^5\,cm^{-2}$).

Kujofsa and Ayers[5] studied combinations of buffer layer thickness and overshoot in ZnS_ySe_{1-y}/GaAs(001) heterostructures utilizing uniform and linearly graded buffer layers that could enable a surface dislocation density on the order of $10^5\,cm^{-2}$. This work also involved the study of the allowable variation in the sulfur mole fraction that would allow the reduction of the threading dislocation density to particular values of interest; for this purpose, threading densities of 3×10^6, 10^7, and $3 \times 10^7\,cm^{-2}$ were considered. The dislocation dynamics model described in Chapter 6 was applied, and the compensation mechanism behavior of overshoot graded structures utilizing a uniform (type A) or linearly graded (type B) buffer layer was investigated. In type A structures, the buffer composition was fixed at y_0. In type B structures, the composition of the linearly graded buffer layer was varied from y_0 at the substrate interface to y_{Buf} at the graded–uniform (buffer–device) layer interface. In both types of structures, the composition in the top (device) layer was y_{Dev}. The thicknesses of the buffer and device layer were h_{Buf} and h_{Dev}, respectively, and the total thickness of the structure was therefore $h = h_{Buf} + h_{Dev}$. For a sulfur mole fraction greater than ~6% at the substrate–buffer interface, misfit dislocations of *negative sense* are introduced to relax the tensile material above. If the composition of the buffer layer is made to overshoot that of the device layer ($y_{Buf} > y_{Dev}$), then misfit dislocations of *positive sense* are introduced to bend over threading dislocations associated with negative-sense misfit dislocations. In the case of high interfacial overshoots, all of the existing mobile dislocations will be bent over at the buffer–device interface and additional half loops will be introduced.

Figures 8.1 through 8.3 illustrate contour plots of the buffer layer thickness as a function of the buffer–device interfacial sulfur composition overshoot for the achievement of various mobile threading dislocation densities using type A and B heterostructures for varying device layer thickness and composition. The results of Figures 8.1 through Figure 8.3 indicate that with the layer thicknesses fixed, there is an amount of compositional overshoot

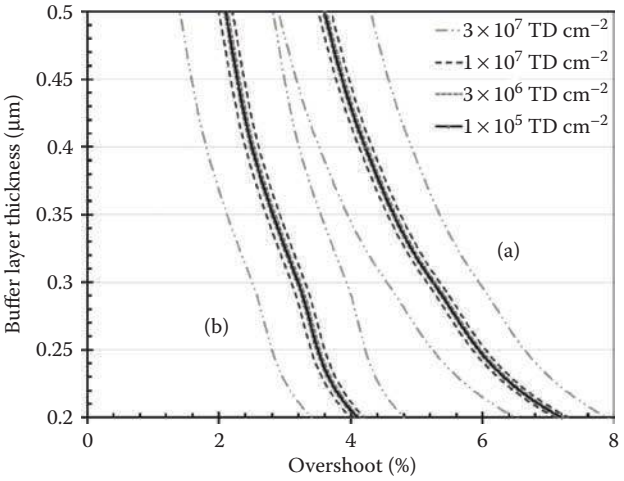


FIGURE 8.1
Contour plot of sulfur composition for the achievement of various mobile threading dislocation densities using (a) a linearly graded ZnS_ySe_{1-y} buffer (type B structure) and (b) an ungraded ZnS_ySe_{1-y} buffer (type A structure) with a device layer thickness of 0.25 μm and a sulfur concentration of 8%. The starting sulfur composition of the buffer layer was fixed at 8%. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 42, 3034 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

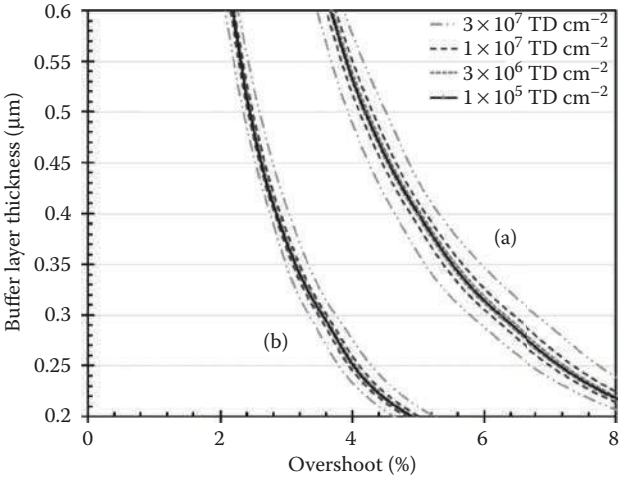


FIGURE 8.2
Contour plot of sulfur composition for the achievement of various mobile threading dislocation densities using (a) a linearly graded ZnS_ySe_{1-y} buffer (type B structure) and (b) an ungraded ZnS_ySe_{1-y} buffer (type A structure) with a device layer thickness of 0.5 μm and a sulfur concentration of 8%. The starting sulfur composition of the buffer layer was fixed at 8%. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 42, 3034 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

at the buffer–device interface that can remove all the mobile threading dislocations. For the practical application of this work, tolerances in the sulfur compositional overshoot were determined for the attainment of dislocation densities equal to 3×10^6, 1×10^7, and 3×10^7 cm^{-2}. The results of the three figures show that the horizontal separation of the contour bands associated with a particular dislocation density diminishes with the target

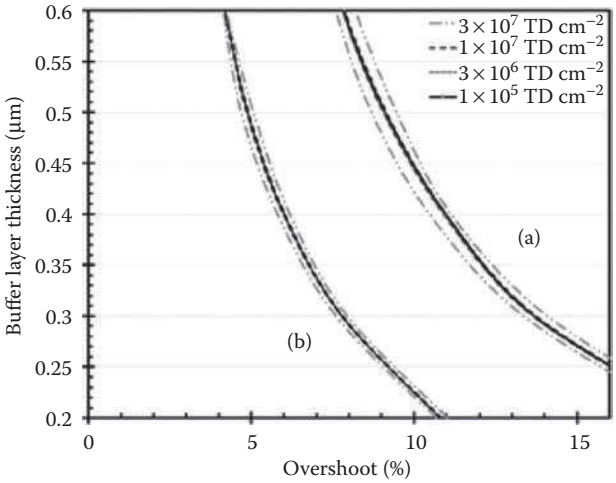


FIGURE 8.3
Contour plot of sulfur composition for the achievement of various mobile threading dislocation densities using (a) a linearly graded ZnS_ySe_{1-y} buffer (type B structure) and (b) an ungraded ZnS_ySe_{1-y} buffer (type A structure) with a device layer thickness of 0.5 μm and a sulfur concentration of 10%. The starting sulfur composition of the buffer layer was fixed at 8%. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 42, 3034 [2013]. With permission. Copyright 2013, Minerals, Metals and Materials Society.)

dislocation density, so that the achievement of 10^6 cm^{-2} or fewer mobile defects will require relatively precise compositional control.

Figure 8.1 compares the contour plots for heterostructures that utilize a linearly graded and ungraded buffer metamorphic buffer layer. The heterostructures modeled in Figure 8.1 have a device layer with 0.25 μm thickness and a sulfur mole fraction of 8%. The starting sulfur composition in the linearly graded buffer layer was fixed at 8%. The optimum buffer layer thickness decreases rapidly and monotonically for increasing overshoot amount. When comparing both types of heterostructures, it becomes clear that for a given buffer layer thickness, the linearly graded buffer requires a higher buffer–device interfacial overshoot to bend over threading dislocations. However, the difference between the average composition ($y_{average} = (y_{Buf} + y_0)/2$) of the buffer and device layers is less in heterostructures with the linearly graded buffer; this may be because the linearly graded buffers allow for the distribution of misfit dislocations throughout the layer, which results in the longest possible misfit dislocation length, and therefore the threading dislocation density is reduced significantly from enhancement of glide motion.

Figure 8.2 shows the contour profile for structures that are similar to those of Figure 8.1 but have a device layer thickness of 0.5 μm. An increase in the device layer thickness results in a more gradual decrease of the buffer layer thickness versus the interfacial overshoot curve. There are two competing mechanisms involved in this behavior: the increase of the device layer thickness increases the lattice relaxation and threading dislocation density in the buffer layer, but at the same time, increasing the device layer thickness enhances the annihilation and coalescence reaction in this layer. Of these, the former appears to be dominant.

Figure 8.3 shows that structures with higher sulfur content exhibit rapid relaxation rates that lead to contraction of the horizontal separation of the contour lines. In this figure, the sulfur mole fraction of the device layer has been increased to $y_{Dev} = 10\%$. For this case, the application of dislocation compensation with a uniform buffer layer requires extremely precise compositional control.

Similar principles should apply to other material systems, but more theoretical and experimental studies are necessary to fully understand the dislocation compensation mechanism.

8.2.2 Temperature-Graded Layers

Kujofsa and Ayers[6] utilized the dislocation dynamics model described in Chapter 6 and investigated threading dislocations in temperature-graded structures with varying thermal budgets. Figure 8.4 shows the average threading dislocation density as a function of the epilayer thickness with initial temperature as a parameter for three different types of temperature profiles. The idea behind this work was to be able to control the glide of threading dislocations and therefore the amount of lattice relaxation by design of the temperature grading profile. Three types of temperature grading profiles were used. Heterostructures with convex-up (type A), linear (type B), and convex-down (type C) temperature grading profiles were considered. For type A structures, the first half of the total epilayer thickness was grown at constant temperature T_0, and in the remaining half of the total thickness, the temperature was linearly decreased from T_0 to a final temperature of $T_F = 300°C$. In linear type grading, the temperature was varied linearly from T_0 to $T_F = 300°C$ during the epitaxial growth of the structure. For type C structures, in the growth of the first 50% of total epilayer thickness, the temperature was linearly decreased from T_0 to a temperature of $T_F = 300°C$, and the remaining growth was continued at this same temperature. The threading dislocation densities in these structures are shown and compared in Figure 8.4. Two-regime behavior is seen in all three cases. Initially, there is a buildup of threading dislocations in the early stages of relaxation, but eventually there is a falloff in the dislocation density due to annihilation and coalescence reactions. Generally, a lower starting temperature for the grade results in a lower threading dislocation density at the surface, but the three types of temperature grading exhibit different behaviors. The convex-down structure exhibits significantly lower threading dislocation densities for thicknesses less than 1 micron, due to the reduced thermal budget. Therefore, design of the temperature grading profile appears to allow control of the threading dislocation density.

8.2.3 Superlattice Buffer Layers

In a number of heteroepitaxial material systems, it has been reported that the insertion of a strained layer superlattice (SLS) can reduce the threading dislocation density of the heteroepitaxial material, compared with the case of direct growth without an SLS. Experimental evidence obtained by cross-sectional transmission electron microscopy (TEM) and etch pit density (EPD) characterization shows that some of the threading dislocations can bend over at the interfaces of the SLS. These may reach the edge of the sample and thereby be removed. However, dislocations that do not reach the wafer edge may serpentine back and forth in the alternating layers of the SLS, which have mismatch strains of opposite sign. Even though these dislocations may not reach the sample edge, they will have increased opportunity to participate in annihilation or coalescence reactions with other dislocations, thus reducing the dislocation density in the overlying material. The efficacy of the edge removal mechanism is expected to reduce with increasing wafer diameter. This is due to the finite glide velocities for dislocations, and also the possibility of pinning or blocking or dislocation motion by other dislocations or other types of crystal defects. On the other hand, the enhancement of the coalescence or annihilation reactions can operate on wafers of arbitrary diameter. Sometimes this mechanism is referred to as dislocation filtering.[7]

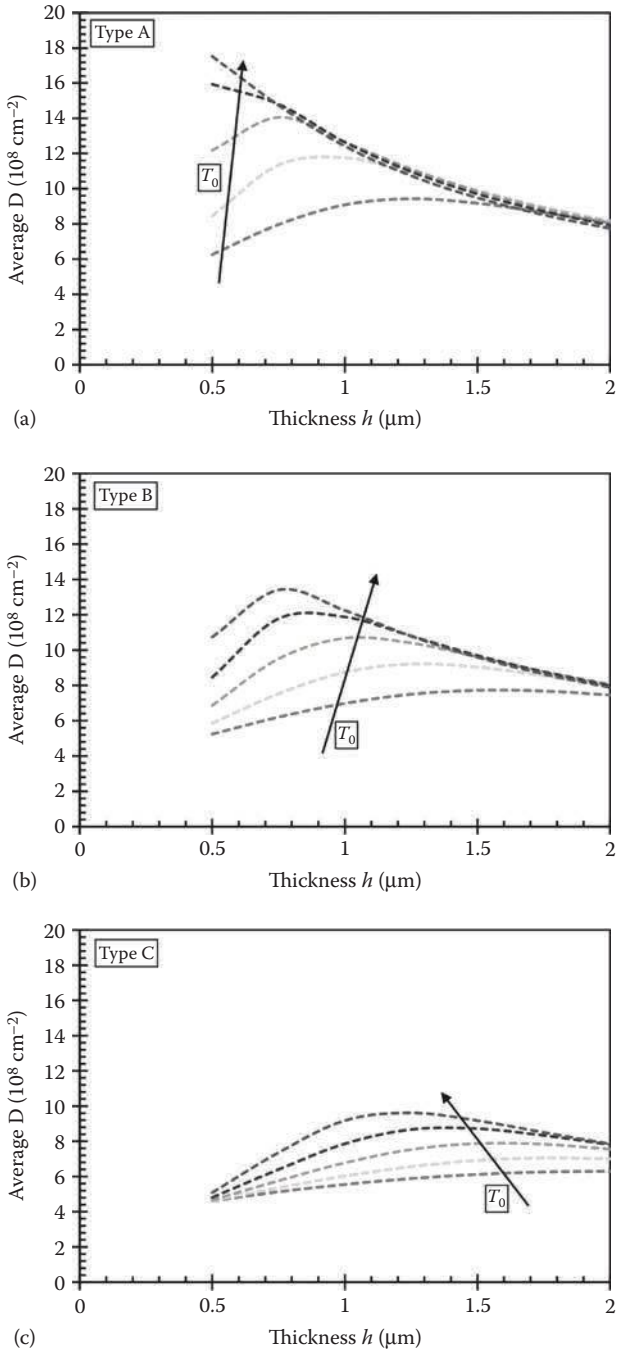


FIGURE 8.4
Average threading dislocation density as a function of the final epilayer thickness h with initial temperature as a parameter for type (a) A, (b) B, and (c) C structures. The initial temperature T_0 is varied from 400°C to 600°C in steps of 50°C and the final T_F temperature is fixed at 300°C. (Reprinted from T. Kujofsa and J. E. Ayers, *J. Electron. Mater.*, 44, 3030 [2015]. With permission. Copyright 2015, Minerals, Metals and Materials Society.)

Soga et al.[8,9] used SLS buffers to produce GaAs-on-Si substrates with reduced threading dislocations compared with direct GaAs/Si(001) heteroepitaxy. In their work, they used metalorganic vapor phase epitaxy (MOVPE) to grow a GaP buffer on Si(001), followed by GaP/$GaAs_{0.5}P_{0.5}$ and $GaAs_{0.5}P_{0.5}$/GaAs superlattices, and finally a thick layer of GaAs. In one set of experiments, they found that GaAs grown on Si(001) with a $GaAs_{0.5}P_{0.5}$/GaAs SLS buffer exhibited an order of magnitude higher photoluminescence (PL) emission intensity than GaAs grown under the same conditions but on a Ge-coated Si substrate. In another set of experiments, molten KOH etching revealed a remarkably low EPD of $4 \times 10^3\ cm^{-2}$, compared with $10^8\ cm^{-2}$, which is typical for direct GaAs/Si(001) heteroepitaxy. It is well established that molten KOH EPDs sometimes underestimate the true threading dislocation density in GaAs on Si; nonetheless, these results suggested a significant reduction in the actual threading dislocation density.

Other workers have reported significant reductions in the dislocation density for GaAs on Si(001) by using SLS dislocation filters. For example, Okamoto et al.[10] used a $Ga_{0.9}In_{0.1}As$/GaAs SLS buffer to reduce to the threading dislocation density in GaAs grown on a Si substrate by MOVPE. The EPD obtained by KOH etching for a 3.5 μm thick layer was $1.4 \times 10^6\ cm^{-2}$, or about two orders of magnitude better than for direct heteroepitaxy.

One novel application of SLS dislocation filters to the growth of GaAs on Si(001) involved the use of $(GaAs)_{1-x}(Si_2)_x$/GaAs superlattices, as reported by Rao et al.[11] In their work, molecular beam epitaxy (MBE) was used to grow GaAs on a vicinal Si(001) substrate with 3 SLSs, each comprising five periods of $(GaAs)_{0.8}(Si_2)_{0.2}$/GaAs. The $(GaAs)_{1-x}(Si_2)_x$ material was grown by migration-enhanced epitaxy (MEE).[12,13] Cross-sectional TEM micrographs showed that all visible dislocations had been filtered by the three SLSs, so that the upper bound for the threading dislocation density in the upper layer of GaAs was $5 \times 10^5\ cm^{-2}$, representing a reduction of the dislocation density by several orders of magnitude.

Rao et al.[11] also found that multiple SLS buffers could be more effective than a single SLS in filtering dislocations. In their structures, with three SLSs of five periods each, there was visible deflection of dislocations at each SLS. The middle SLS was most effective, however, affecting a two-order reduction in the dislocation density. It is likely that SLS buffers are less effective at filtering dislocations when their density is either very high (due to the limited mismatch) or very low (limited by the available dislocations).

Qualitatively similar results have been obtained in other material systems, such as GaSb on GaAs(001). Qian et al.[14] investigated the use of SLS buffers to reduce the threading dislocation density in this material system using MBE. They found that the insertion of a five-period GaSb/AlSb SLS reduced the threading dislocation in a 1.1 μm thick GaSb layer by more than an order of magnitude compared with the case of direct heteroepitaxy. Here, the threading dislocation densities were characterized using plan-view TEM. The dislocation filtering action of the SLS is shown in Figure 8.5. In this bright-field transmission electron micrograph, it can be seen that most of the threading dislocations are bent over in the SLS, resulting in a low dislocation density in the top GaSb layer.

SLS dislocation filters have also been applied to the growth of II-VI semiconductors on mismatched substrates. For example, Reno et al.[15] used SLS buffers in the growth of $Cd_{0.955}Zn_{0.045}Te$ on GaAs(001) by MBE. The lattice mismatch strain in their system is $f = -13.7\%$. The samples investigated were grown on GaAs(001) substrates, starting with 30 Å of ZnTe (to establish the (001) orientation), a 3 μm thick $Cd_{0.955}Zn_{0.045}Te$ buffer, an SLS buffer, and finally, a top layer of $Cd_{0.955}Zn_{0.045}Te$. The SLS comprised 15 periods of $Cd_{0.91}Zn_{0.09}Te$/CdTe. It was designed to have an average lattice constant matching the top $Cd_{0.955}Zn_{0.045}Te$ epitaxial layer, and the period of the SLS was varied in order to find the optimum value. They found that the SLS was most effective in filtering dislocations

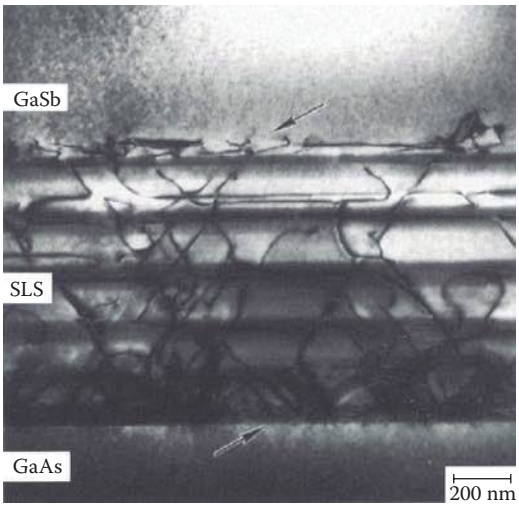


FIGURE 8.5
Bright-field cross-sectional TEM image showing the dislocation filtering action of a GaSb/AlSb SLS inserted between a GaAs(001) substrate and the GaSb top layer grown by MBE. (Reprinted from W. Qian et al., *J. Electrochem. Soc.*, 144, 1430 [1997]. With permission. Copyright 1997, Electrochemical Society.)

when its period was approximately 2600 Å. In these structures, bright-field cross-sectional TEM images revealed that the $Cd_{0.955}Zn_{0.045}Te$ buffer had a threading dislocation density of 10^{10}–10^{11} cm^{-2}, but that the layer above the SLS had a much reduced threading dislocation density of $<10^5$ cm^{-2}. The average threading dislocation density may have been higher, based on their minimum 004 x-ray rocking curve width of 175 arcsec. Nonetheless, the SLS buffer affected a reduction in the dislocation density by orders of magnitude.

Early efforts to apply SLS dislocation filters to GaN on sapphire (0001) were unfruitful. Using MOVPE, Qian et al.[16] inserted three periods of 6 cm GaN/6 cm AlN on a 0.5 μm GaN layer grown on sapphire (0001), and then grew an additional 3.0 μm of GaN. Cross-sectional and plan-view TEM investigation showed no evidence of dislocation bending at the SLS. However, Qian et al. noted that the threading dislocations in GaN are on $\{1\bar{1}00\}$-type slip planes. Therefore, the stresses of the SLS in the (0001) plane would not provide a driving force for glide of dislocations on their $\{1\bar{1}00\}$ slip planes.

Other efforts to apply SLS dislocation filters to GaN-on-silicon and sapphire substrates have met with limited success. For example, Feltin et al.[17] studied the dislocation filtering properties of GaN/AlN SLSs in the MOVPE growth of GaN on Si(111). They found that the threading dislocation density could be reduced by the insertion of SLSs; however, the reduction was only by a factor of ¼ (from 1.6×10^{10} to 4×10^9 cm^{-2}) when four SLSs were inserted.

Sun et al.[18] studied the use of $AlN/Al_{0.85}Ga_{0.15}N$ SLSs grown by pulsed atomic layer epitaxy (PALE) to filter dislocations in ~1.0 μm thick $Al_{0.55}Ga_{0.45}N$ layers grown on sapphire (0001). Based on cross-sectional TEM imaging, they found that the SLS could effectively block screw dislocations, but had a negligible effect on edge-type dislocations. A comparison between the SLS sample and a control sample with no SLS showed that the density of screw-type threading dislocations was reduced by more than an order of magnitude, from 4×10^9 to 3×10^8 cm^{-2}. Of course, the effect on the overall threading dislocation density was less dramatic.

Gourley et al.[7] suggested that semiconductor superlattices might filter threading dislocations even without built-in strains, provided that the individual layers in the superlattice differ in elastic stiffness. To test this concept, they carried out a set of experiments involving the growth of $In_xGa_{1-x}As$ on GaAs(001) substrates by MBE. In all experiments, an $In_zGa_{1-z}As$ buffer was first grown on the GaAs(001) substrate, followed by a $In_xGa_{1-x}As/In_yAl_{1-y}As$ superlattice, and then a top layer of $In_zGa_{1-z}As$. The individual layers in the superlattice were 10 nm thick. Two cases were investigated. For structures having $x = y = z$ (case I), the superlattice was lattice matched with the buffer layer and top layer, and therefore expected to be essentially strain-free. For structures with $x \neq y$ and $z = (x + y)/2$ (case II), the superlattice layers had equal and opposite strains. The strains, compositions, and superlattice period were determined by x-ray diffraction (XRD). The dislocations in the structures were imaged using photoluminescence microscopy (PLM). Two different excitation wavelengths (647.1 and 406.7 nm, "deep" excitation and "shallow" excitation, respectively) were used to probe different depths of the samples. It was found that both the lattice-matched and strained superlattices were effective in filtering dislocations, with a significant (at least an order of magnitude) reduction of the dislocation density in the top layer.

El-Masry et al.[19] studied the mechanisms for filtering of dislocations in GaAs on Si(001) by SLSs. In this work, $GaAs_{1-y}P_y/In_xGa_{1-x}As$ SLSs ($y = 2x$) lattice matched to GaAs were used, and cross-sectional TEM characterization was used to study the resulting dislocation interactions. They described five different experimentally observed interactions between dislocations and the SLS, shown schematically in Figure 8.6. As shown in Figure 8.6a, an edge dislocation may experience a zero Peach–Koehler force and therefore not bend at

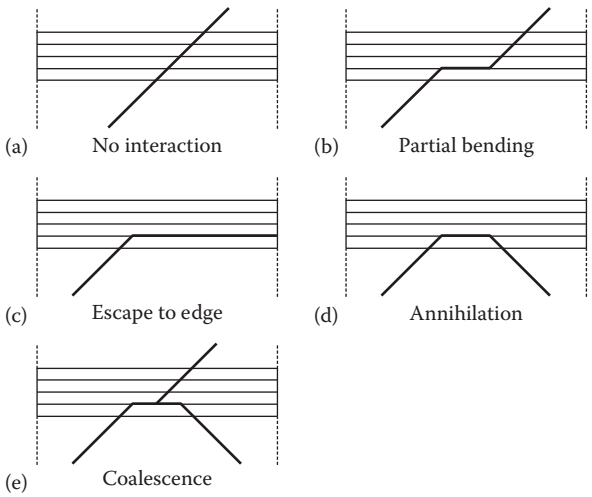


FIGURE 8.6
Five types of interactions between dislocations and an SLS buffer, as experimentally observed by El-Masry et al.[19] (a) An edge dislocation may experience a zero Peach–Koehler force and therefore not bend at the SLS. (b) Some dislocations may jog at the SLS without being removed from the top layer. (c) A mixed dislocation may bend over and glide all the way to the edge, resulting in the elimination of a threading dislocation in the top layer. (d) Dislocations of opposite Burgers vector may participate in an annihilation reaction, whereby a half loop is created but two threading dislocations are eliminated from the top layer. (e) Two dislocations may coalesce to form a third dislocation. This process removes one threading dislocation from the upper layer.

the SLS. In other cases, such as an edge dislocation on a vicinal substrate, there may be an insufficient Peach–Koehler force on the dislocation; it may therefore jog in the SLS as in Figure 8.6b, but there will be no reduction of the dislocation density. On the other hand, a mixed dislocation such as that shown in Figure 8.6c can bend over completely and be removed from the upper epitaxial layer, as long as it can glide all the way to the sample edge. Dislocations of opposite Burgers vector may react as shown in Figure 8.6d, with the resultant removal of two threading dislocations from the upper layer. Finally, two dislocations may react to form a third dislocation, as in Figure 8.6e. This process removes one threading dislocation from the upper layer.

Based on the available experimental evidence, it appears that superlattice buffer layers are generally applicable for the reduction of threading dislocation densities in mismatched heteroepitaxial zinc blende semiconductors. Both SLSs and superlattices with modulated elastic stiffness are effective as dislocation filters. In the former case, the strains in the layers can cause dislocations to weave back and forth, thus promoting annihilation and coalescence reactions between dislocations. In superlattices with modulated elastic stiffness, dislocations will tend to bend over into the softer material with the same end result. A single dislocation filter will be most effective at a moderate threading dislocation density, but less effective at both lower and higher dislocation densities. However, in materials with high threading dislocation densities, the effectiveness of the dislocation filtering can be enhanced by using multiple superlattices.

There is also some evidence to guide the design of superlattice dislocation filters. If the strain filter mechanism is to be used, then the SLS must be designed to promote the bending over of existing dislocations without the introduction of new ones. Thus, the individual layers of the SLS must grow in a Frank–van der Merwe (layer-by-layer) mode, and they must exceed the Matthews and Blakeslee critical layer thickness for the bending over of grown-in dislocations. However, they must not have sufficient strain or thickness to promote significant dislocation multiplication or nucleation. The overall SLS stack must not exceed its Matthews and Blakeslee critical layer thickness, for this would introduce new dislocations. However, this problem can usually be avoided by designing the SLS to be strain balanced. (The alternate layers will have equal but opposite strains built in.) In practical SLS dislocation filters, the individual layers of the SLS should have moderate strain and thicknesses that are greater than the Matthews and Blakeslee critical layer thickness, but not by more than about a factor of 10. Qian et al.[14] found that a GaSb/AlSb superlattice was most effective in reducing the dislocation density for a GaSb layer grown on a GaAs(001) substrate when the GaSb and AlSb layers were both about 100 nm thick. The Matthews and Blakeslee critical layer thickness for GaSb on a thick AlSb substrate (or AlSb on a thick GaSb substrate) is about 50 nm, so the thicknesses in the optimum SLS were about twice h_c. In the work of Reno et al., the optimum period for a $Cd_{0.91}Zn_{0.09}Te/CdTe$ SLS (15 periods) to filter dislocations in a $Cd_{0.955}Zn_{0.045}Te$ layer grown on a GaAs(001) substrate was 260 nm. Therefore, the individual layers were 130 nm thick, or about six times the Matthews and Blakeslee critical layer thickness for $Cd_{0.91}Zn_{0.09}Te/CdTe$ ($|f| = 0.55\%$ and $h_c = 20$ nm).

The limited success of SLS buffers for the dislocation filtering in (0001) nitride materials appears to be related to the geometry of the slip systems and specifically the inability to filter edge-type threading dislocations. It is possible, however, that SLS dislocation filters may prove to be more effective when used with semipolar or nonpolar orientations or when combined with other defect engineering techniques.

Despite the large body of experimental work, there remains a need for a general quantitative model that can be used to design SLS dislocation filters.

8.3 Reduced Area Growth Using Patterned Substrates

In heteroepitaxial layers with moderate mismatch ($|f| < 2\%$), the initial growth is pseudomorphic and the initiation of lattice relaxation occurs by the bending over of substrate dislocations (the Matthews and Blakeslee mechanism). However, the number of available threading dislocations N to participate in this lattice relaxation process depends on the growth area according to $N = DA$, where D is the density of substrate threading dislocations (which are replicated in the epitaxial layer) and A is the growth area. Therefore, a reduction in the growth area can also reduce the density of misfit dislocations at the interface.[20] In fact, if the growth area is reduced sufficiently, there may be no substrate threading dislocations available to participate in lattice relaxation. Other mechanisms involving dislocation nucleation can become active, but only at thicknesses much greater than the critical layer thickness for the Matthews and Blakeslee mechanism. Therefore, a reduction in the growth area may even enable the achievement of metastable heteroepitaxial layers that are completely free from both misfit and threading dislocations, even though they are greater than the critical layer thickness. Even if sources of heterogeneous dislocation nucleation become active, they are also expected to have a finite density so that reduced growth area can suppress relaxation by these as well.

To test these ideas, Fitzgerald et al.[21,22] performed a series of experiments involving the MBE growth of $In_{0.05}Ga_{0.95}As$ layers on GaAs(001) substrates that had been patterned with round or rectangular 2 μm high mesas. Experiments were conducted with epitaxial layer thicknesses of 350, 700, and 825 nm, all many times the expected critical layer thickness. Various substrates were utilized, with threading dislocation densities varying from 10^2 to $1.5 \times 10^5\ cm^{-2}$.

Figure 8.7 shows cathodoluminescence (CL) images of 350 nm thick $In_{0.05}Ga_{0.95}As$ grown on a GaAs(001) substrate with a dislocation density of $1.5 \times 10^5\ cm^{-2}$. The layer of Figure 8.7a

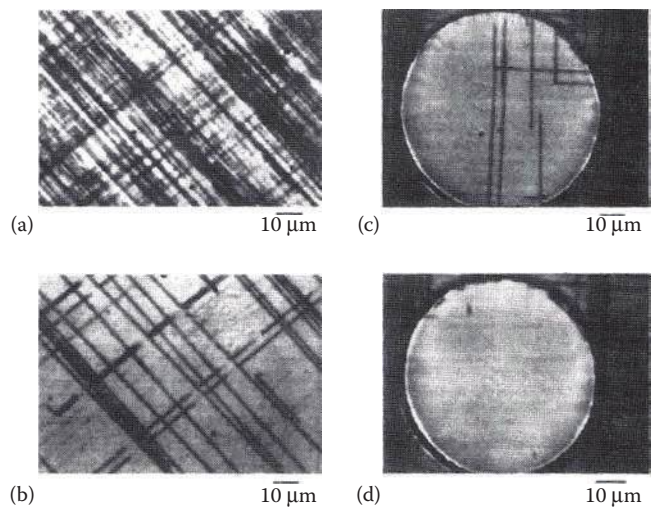


FIGURE 8.7
Cathodoluminescence images of 350 cm thick $In_{0.05}Ga_{0.95}As$ grown on a GaAs(001) substrate with a dislocation density of $1.5 \times 10^5\ cm^{-2}$: (a) large-area growth, (b) growth on a 200 μm circular mesa, (c) growth on a 90 μm circular mesa, and (d) growth on a 67 μm circular mesa. (Reprinted from E. A. Fitzgerald et al., *J. Appl. Phys.*, 65, 2220 [1989]. With permission. Copyright 1989, American Institute of Physics.)

was grown over a large area, without mesa patterning. The samples of Figure 8.7b–d were grown on circular mesas having diameters of 200, 90, and 67 μm, respectively. It can be clearly seen that the misfit dislocation density decreases with the size of the mesa. It is also evident that there are different numbers of dislocations along the two perpendicular [110] directions, for the reduced area growth.

A qualitative analysis revealed that for the 350 nm thick $In_{0.05}Ga_{0.95}As$ layers, the linear misfit dislocation density increased linearly with the mesa diameter. And, as expected, higher misfit dislocation densities were measured on substrate having higher threading dislocation densities. The results are shown in Figure 8.8.

The data of Figure 8.8a correspond to a substrate threading dislocation density of $1.5 \times 10^5\ cm^{-2}$. For both the α misfit dislocations running along the $[1\bar{1}0]$ direction and the β dislocations running along the [110] direction, the average misfit dislocation density (in cm^{-1}) increased linearly with the mesa diameter. Also, both dislocation densities extrapolate to zero at zero mesa diameter; this indicates that only area-dependent dislocation sources are active, and eliminates the possibility of dislocation multiplication in these samples. However, there is a marked difference between the densities for the two types of dislocations, and the α dislocations appear to nucleate roughly twice as much as the β dislocations.

For the samples of Figure 8.8b, the substrate dislocation density was $10^4\ cm^{-2}$. The trends are qualitatively the same as in Figure 8.8a. The average linear densities of misfit dislocations are lower than those for the case of the more dislocated substrate, but the improvement is not as much as might be expected if only the substrate dislocations are active as sources of misfit dislocations. This implies the existence of other fixed sources that scale with the area.

This behavior is even more evident in Figure 8.8c, for which the substrate threading dislocation was $10^2\ cm^{-2}$. Here, the misfit dislocation densities are similar to the case shown in Figure 8.8b, even though the substrate threading dislocation has been reduced by two orders of magnitude. Therefore, there are significant fixed sources of misfit dislocations in addition to the substrate threading dislocations. However, the misfit dislocation densities still extrapolate to zero at zero mesa area, so a reduction in the growth area can dramatically decrease the misfit dislocation density at the interface.

Fitzgerald et al. also investigated 700 nm thick layers of $In_{0.05}Ga_{0.95}As$ on GaAs(001) substrates that had been mesa patterned. For these thicker layers, the average linear misfit dislocation densities increased super linearly with the mesa diameter, for mesas larger than 200 μm. This shows that dislocation multiplication was active in the thicker layers with large mesa diameters. Also, for mesas smaller than 200 μm, the misfit dislocation densities did not extrapolate to zero for zero mesa diameter. This indicates that there are dislocation sources that do not scale with the mesa size, and these are believed to be associated with the mesa edges.

In conclusion, it has been shown that a reduction in the growth area can reduce the densities of misfit dislocations in mismatched heteroepitaxial layers that are greater than the critical layer thickness. For $In_{0.05}Ga_{0.95}As$ layers that were five times the experimentally determined critical layer thickness, it is possible to grow material entirely free from misfit dislocations if the growth area is reduced sufficiently. Only fixed dislocation sources that scale with the mesa size are active, and no dislocation multiplication occurs in these layers. For $In_{0.05}Ga_{0.95}As$ layers that were 10 times the experimentally determined critical layer thickness, new sources of dislocations came into play. For small mesas, sources associated with the mesa edges, which do not scale with mesa area, become important. For large mesas, dislocation multiplication becomes important. Nonetheless, in all cases the misfit dislocation density can be reduced

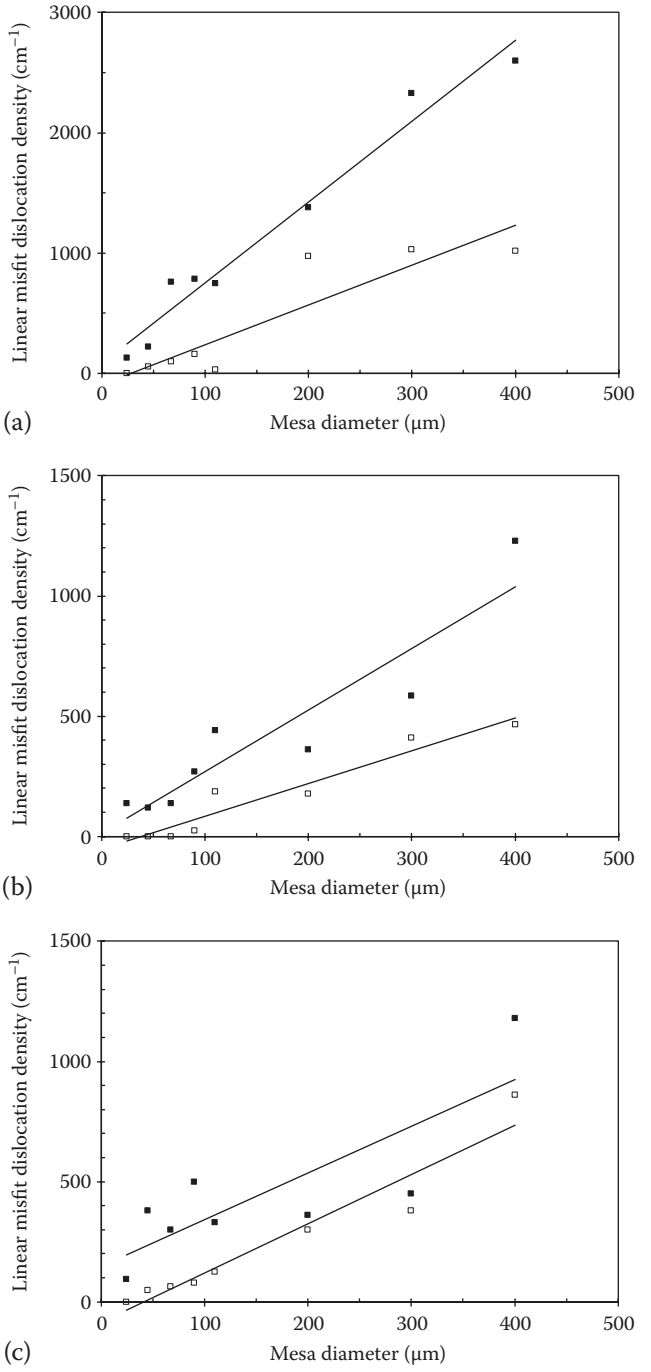


FIGURE 8.8
Average linear densities of interfacial misfit dislocations for 350 nm thick $In_{0.05}Ga_{0.95}As$ layers on patterned GaAs(001) substrates, as functions of the round mesa diameter. The filled squares are for α misfit dislocations running along the [1$\bar{1}$0] directions, whereas the open squares are for β dislocations running along the [110] directions. The substrate threading dislocation densities were (a) $1.5 \times 10^5\,cm^{-2}$, (b) $10^4\,cm^{-2}$, and (c) $10^2\,cm^{-2}$. (Reprinted from E. A. Fitzgerald et al., *J. Appl. Phys.*, 65, 2220 [1989]. With permission. Copyright 1989, American Institute of Physics.)

by a reduction in the growth area. Qualitatively similar results should be expected for other material systems, and this has been reported in the case of SiGe/Si with high Ge content.[23] Therefore, reduced growth area should be useful for the elimination of interfacial defects in heterojunction devices, if layers thicker than the critical layer thickness are required.

8.4 Patterning and Annealing

It was shown Section 8.3 that reduced area epitaxy can reduce, or even eliminate, the introduction of misfit dislocations at the interface of a mismatched heteroepitaxial layer. That approach can extend the usable layer thickness for heterojunction devices up to perhaps 10 times the critical layer thickness. But many heteroepitaxial systems of interest exhibit high mismatch and a Volmer–Weber growth mode, so it is not possible to obtain pseudomorphic layers by a reduction in the growth area.

Instead, however, we can remove threading dislocations from the material after it has relaxed through the use of patterning and annealing. This approach, proposed by Zhang et al.,[24] is known as PHeP or dislocation sidewall gettering (DSG). The reduced lateral dimensions of the epitaxial material allow threading dislocations to glide to the sidewalls, where they are removed. If the pattern dimensions are small enough, sidewall image forces will attract the threading dislocations and effectively getter them.

There are two embodiments of the PHeP approach. In the first, a planar heteroepitaxial layer is grown, then patterned by etching, and finally annealed at an elevated temperature. In the second version of PHeP, the substrate is patterned prior to growth. Either mesa patterning can be used or an oxide mask layer can be used with selective epitaxy. If the substrate is patterned prior to growth, the annealing can occur during the growth itself, or during a postgrowth anneal.

Zhang et al. reported a simple model for PHeP,[25] which can be summarized as follows. For a relaxed heteroepitaxial layer, much greater than the critical layer thickness, the linear density of misfit dislocations ρ_{MD} is

$$\rho_{MD} = \frac{f}{b \sin\alpha \sin\phi}\left(1 - \frac{h_c}{h}\right), \tag{8.1}$$

where:

f	is the lattice mismatch
α	is the angle between the Burgers vector and line vector for the misfit dislocations
ϕ	is the angle between the interface and the normal to the slip plane
$b\sin\alpha\sin\phi$	is the misfit-relieving component of the Burgers vector
h	is the layer thickness
h_c	is the critical layer thickness

The misfit dislocation density and the threading dislocation density D are related by the mean length of the misfit dislocation segments, L_{MD}. If there are two orthogonal misfit dislocation arrays with the same value of L_{MD} and each misfit dislocation has n threading segments associated with it, then

$$\rho_{MD} = \frac{DL_{MD}}{2n}. \tag{8.2}$$

Here, n can range from 0 to 2, $n = 1$ corresponds to the substrate associated dislocations and for dislocation half loops, $n = 2$.

Zhang assumed that for an unpatterned (planar) heteroepitaxial layer, much greater than the critical layer thickness, dislocation multiplication processes would have created plentiful threading dislocations, and so any additional lattice relaxation would occur by their glide rather than the creation of new threading dislocations. Then, if the threading dislocation density is fixed and the equilibrium strain is maintained,

$$L_{MD} = \frac{f}{b\sin\alpha\sin\phi}\left(1-\frac{h_c}{h}\right)\frac{2n}{D}. \tag{8.3}$$

For the planar layer, it can be assumed that n and D will remain unchanged, as long as threading dislocations do not move long enough distances to encounter an edge, and if there is negligible threading dislocation annihilation.

On the other hand, threading dislocations may reach the edges if the heteroepitaxial material is patterned into mesas, leading to a reduction in n and D. Suppose the mesas are square, with sides of length L parallel to the ⟨110⟩ directions in the (001) interface of a zinc blende or diamond semiconductor. Threading dislocations located within a distance Δ from a sidewall can be removed by glide under the influence of the image forces. This leads to a decrease in the average value of n, and therefore D, during thermal processing of such patterned layers. For a threading segment located at a distance r (along a (111) glide plane) from a sidewall, as shown in Figure 8.9, the attractive image force[26] is approximately

$$F_I = \frac{Gb^2h}{4\pi r\cos\lambda}\left(\cos\alpha+\frac{\sin\alpha}{(1-\nu)}\right), \tag{8.4}$$

where:

- G is the shear modulus
- λ is the angle between the threading segments and the interface
- α is the angle between the Burgers vector and the line vector for the threading segment
- ν is the Poisson ratio

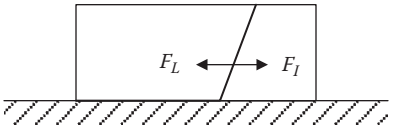


FIGURE 8.9
Removal of a threading dislocation from a patterned heteroepitaxial layer under the influence of the image force. F_L is the line tension in the misfit segment and F_I is the attractive image force associated with the mesa sidewall.

Neglecting the Peierls forces on the threading dislocation, the image force is opposed by the line tension in the misfit segment, which is given approximately by

$$F_L = \frac{Gb^2}{4\pi}\left(\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right)\ln\left(\frac{R}{2b}\right), \tag{8.5}$$

where R is one-half the spacing between dislocations (perpendicular to the intersection of the glide plane and the interface) or the layer thickness, whichever is smaller. In the case of nearly complete relaxation ($h >> h_c$), the spacing of the misfit dislocations is approximately $b\sin\alpha\sin\phi/|f|$. The condition for the glide of a dislocation to a sidewall is $F_I > F_L$. If we consider the worst case of a threading dislocation located at the center of the mesa, then $r = L/2$. The condition for removal of the threading dislocation by glide to the sidewall can therefore be written as

$$L < \frac{\left(\frac{2h}{\cos\lambda}\right)\left(\cos\alpha + \frac{\sin\alpha}{(1-\nu)}\right)}{\left(\cos^2\alpha + \frac{\sin^2\alpha}{4(1-\nu)}\right)\ln\left(\frac{\sin\alpha\sin\phi}{4|f|}\right)}. \tag{8.6}$$

It is expected that threading dislocations can be removed from the periphery of a mesa having arbitrary shape, as long as the dislocations are within a distance Δ from the sidewall. Here, Δ can be considered the active range of the image forces and is equal to one-half of the critical value of L calculated above. For the (001) heteroepitaxy of zinc blende or diamond semiconductors, it has been estimated as[25]

$$\Delta = \frac{8h}{\ln\left(1/4|f|\right)}. \tag{8.7}$$

Therefore, neglecting dislocation–dislocation interactions and the Peierls force, threading dislocations can be removed completely from square-patterned regions of size 2Δ.[25] Zhang et al. calculated engineering curves for the application of PHeP, which predict the maximum mesa size for which all threading dislocations may be removed by glide to the sidewalls, and the results are shown in Figure 8.10.

Zhang et al. investigated the application of the PHeP process to ZnSe/GaAs(001) and $ZnSe_{1-x}S_x$/GaAs(001) grown by photoassisted MOVPE. In this work, planar layers were grown, and following growth, some of the layers were patterned and annealed. The threading dislocation densities in the heteroepitaxial material were determined using crystallographic etching (6 s in a 0.4% bromine-in-methanol solution at 300K).

In order to study the basic mechanism of PHeP, one ZnSe/GaAs(001) wafer was grown and cut into pieces that underwent different processes. The epitaxial layer thickness was 600 nm. From this wafer, four types of samples were produced: (1) as grown, (2) postgrowth annealed, (3) postgrowth patterned, and (4) PHeP prepared. The postgrowth annealing was conducted for 30 min at 600°C in flowing hydrogen. Figure 8.11 shows the etch pit morphology of (a) the as-grown layer and (b) the PHeP-prepared material cut from the same wafer. The as-grown layer had an EPD of 10^7 cm^{-2}, whereas the PHeP material was completely free of threading dislocations even in the largest 70 μm × 70 μm patterned regions. This corresponds to an EPD of less than 2.0×10^4 cm^{-2}, and at least a 500-fold reduction

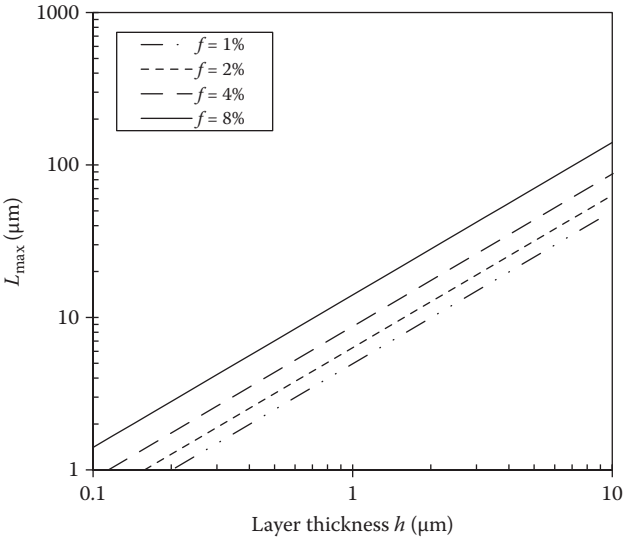


FIGURE 8.10
Dislocation engineering curves for PHeP. L_{max} is the maximum mesa size for which all of the threading dislocations can be removed by glide to the sidewalls. (Adapted from X. G. Zhang et al., *J. Electron. Mater.*, 27, 1248 [1998].)

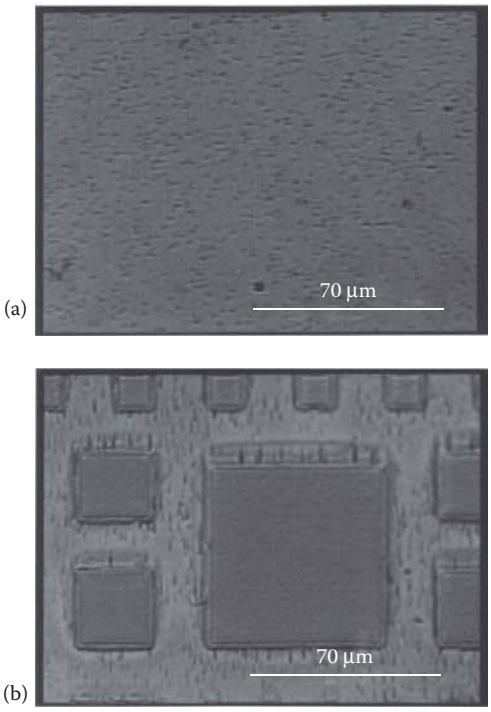


FIGURE 8.11
Etch pit morphology for two 600 cm thick ZnSe/GaAs(001) samples that were processed differently: (a) as grown and (b) patterned and annealed 30 min at 600°C. (Reprinted from X. G. Zhang et al., *J. Appl. Phys.*, 91, 3912 [2002]. With permission. Copyright 2002, American Institute of Physics.)

compared with the as-grown layers. In fact, this value is even lower than that of the GaAs substrate (EPD = 10^5 cm^{-2}). The samples that underwent patterning alone or annealing alone exhibited the same threading dislocation density as the as-grown layer, within experimental error. The result for layers that were annealed without patterning is consistent with the early studies of Chand and Ghu.[27] This indicates that for partially strain-relaxed layers with large lateral dimensions, threading dislocations may not move long enough distances to encounter an edge easily, and annealing alone causes very little threading dislocation annihilation or threading dislocation recombination. Patterning alone has no effect on the threading dislocation density. Therefore, the removal of threading dislocations by PHeP involves thermally activated dislocation motion in the presence of sidewalls.

To investigate the behavior of the PHeP process for different layer thicknesses, a series of ZnSe/GaAs(001) wafers were processed. The layer thicknesses varied from 200 to 1200 nm. Each wafer was cut so that the EPD could be measured for the as-grown layer and also after patterning and annealing. In the case of PHeP-processed wafers, the anneal was conducted for 30 min at 600°C in flowing hydrogen.

The EPDs for as-grown wafers were all of the order of 10^7 cm^{-2}. The nearly constant threading dislocation density may indicate that there is little dislocation annihilation or coalescence. This may be a result of the low growth temperature used for photoassisted MOVPE growth.

For the patterned and annealed wafers, the EPD decreased monotonically with increasing layer thickness, as shown in Figure 8.12 for the case of 70 μm wide mesas. No etch pits were observed in the layers of thickness ≥300 nm. Qualitatively, these results are consistent with the Zhang et al. model. Also, in the case of incomplete etching, it was found that it was the mesa sidewall height, rather than the total epitaxial layer thickness, that determines the effectiveness of PHeP. This is because the lateral forces acting on threading dislocations are proportional to the sidewall height.

To study the effect of annealing temperature on the threading dislocation reduction by PHeP, a set of otherwise identically prepared 300 nm ZnSe/GaAs(001) patterned samples

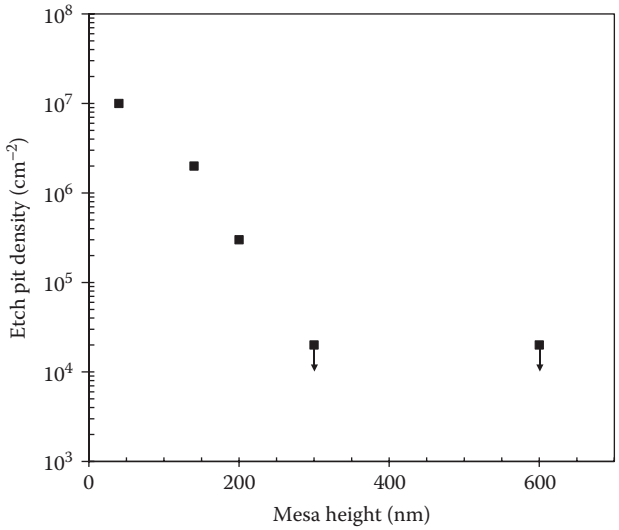


FIGURE 8.12
EPD vs. layer thickness for mesa-etched and annealed ZnSe/GaAs(001). The data shown are for 70 μm square mesas. The annealing was conducted for 30 min at 600°C in flowing hydrogen. (Adapted from X. G. Zhang et al., *J. Appl. Phys.*, 91, 3912 [2002]. With permission. Copyright 2002, American Institute of Physics.)

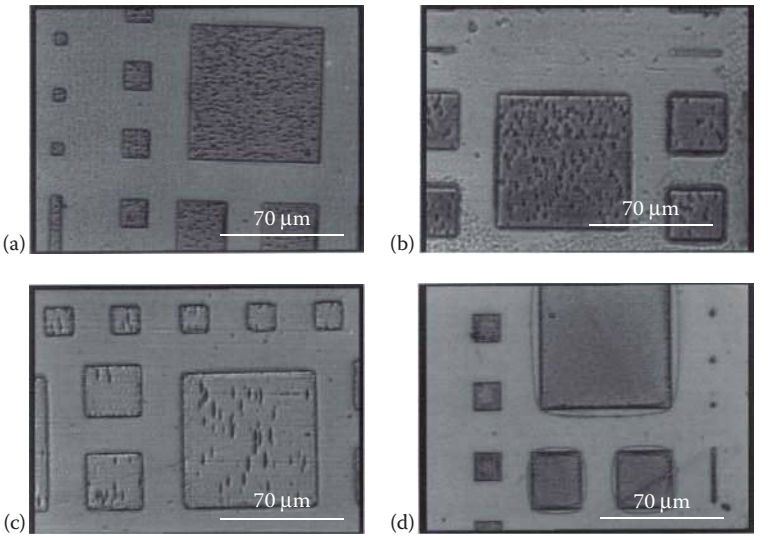


FIGURE 8.13
Etch pit morphology for 300 nm ZnSe/GaAs(001) samples that were patterned and annealed for 30 min at different temperatures: (a) 400°C, (b) 450°C, (c) 475°C, and (d) 500°C. (Reprinted from X. G. Zhang et al., *J. Appl. Phys.*, 91, 3912 [2002]. With permission. Copyright 2002, American Institute of Physics.)

were annealed at different temperatures in the range of 400°C–600°C for 30 min. Figure 8.13 shows the etch pit morphology of layers annealed at 400°C, 450°C, 475°C, and 500°C. The sample annealed at 400°C of Figure 8.13a exhibits the same EPD as the as-grown sample, approximately 10^7 cm^{-2}. Annealing at 450°C (Figure 8.13b) results in a reduction of the EPD to a value of 3.5×10^6 cm^{-2} and at 475°C (Figure 8.13c) to a value of 9×10^5 cm^{-2}. When the annealing temperature is raised to 500°C (Figure 8.13d) or above, PHeP results in complete removal of threading dislocations from 70 µm × 70 µm patterned regions. Zhang et al. plotted the results in the form of dD/dt on an Arrhenius plot and obtained an activation energy of 0.7 eV, which corresponds roughly to the activation energy for dislocation glide, reported to be 1 eV for bulk ZnSe.[28]

The model of Zhang et al. predicts counterintuitively that PHeP should be more effective for heteroepitaxial layers with higher mismatch. Zhang et al. reported a preliminary study of PHeP applied to $ZnS_{0.02}Se_{0.98}$ layers on GaAs, which exhibit approximately two-thirds the lattice mismatch compared with ZnSe on GaAs (+0.18% vs. +0.27%). Both material systems have the same sign of lattice mismatch. They found that, whereas for ZnSe/GaAs(001) all threading dislocations could be removed from mesas having aspect ratios of $W/h < 250$, the dislocations could not be removed completely from $ZnS_{0.02}Se_{0.98}$/GaAs(001) even with a mesa aspect ratio of $W/h = 200$. Further work is necessary, however, to clarify the mismatch dependence.

8.5 Defect Reduction by Selective Evaporation

Ledentsov et al.[29–32] have demonstrated a unique defect reduction technique (DRT) applicable to quantum dots, which involves selective evaporation of defected dots. Their particular

work involved the fabrication of InAs quantum dots on InGaAs buffer layers on GaAs substrates for quantum dot lasers operating at wavelengths in the range of 1.4–1.5 μm for telecommunication applications. They observed that the larger dots tended to be the ones containing dislocations, as would be expected on the basis of strain energy considerations. Therefore, by covering the quantum dots with a partial cap of GaAs, the larger dots were left partially uncapped. Annealing at a high temperature resulted in the evaporation of the incompletely capped dots, which were defected, while leaving the capped dots in place. This was found to eliminate most of the quantum dots containing dislocations.

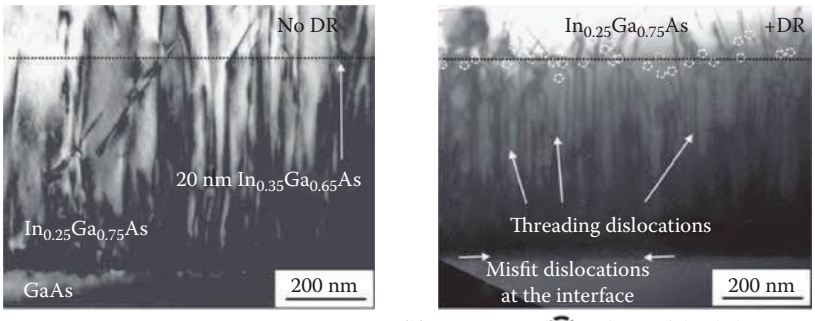
They found that some of the smaller quantum dots were also defected, due to the random nature of the introduction of dislocations. However, capping with AlAs was selective, and any defected dot was left uncapped due to the strain field of the dislocation penetrating the top. The AlAs allowed annealing at an even higher temperature, and it was possible to evaporate even the smaller dots that contained dislocations, so that essentially all defected dots could be eliminated and only the undefected dots remained for laser fabrication.

In subsequent work, Ledentsov et al.[33] showed that this general approach could be extended to continuous layers, in which case the dislocations are bent over so that they do not penetrate the surface. This is illustrated schematically in Figure 8.14. The threading dislocation, shown as a black line, is replicated in an epitaxial layer grown on top. Selective growth of a cap layer leaves the termination of the dislocation uncapped, so evaporation of the uncapped material removes the dislocation from the epitaxial layer. Additional growth of the cap material causes the dislocation to bend over rather than penetrate to the surface.

The application of this approach to $In_{0.25}Ga_{0.75}As$ on an InGaAs metamorphic buffer on a GaAs substrate is shown in the TEM cross-sectional images of Figure 8.15. The image in Figure 8.15a shows a structure grown without the use of the DRT. Figure 8.15b shows a similar structure in which the DRT was used to bend over threading dislocations.



FIGURE 8.14
Generalized DRT involving selective evaporation of material in the strain field of a threading dislocation, followed by overgrowth that results in bending over of the original dislocation. (Reprinted from N. N. Ledentsov et al., *J. Cryst. Growth*, 301–302, 914 [2007]. With permission. Copyright 2007, Elsevier.)



(a) InGaAs QW (b) QW + DRT - tilting of the dislocations

FIGURE 8.15
Cross-sectional TEM images showing the application of the DRT to $In_{0.25}Ga_{0.75}As$ on GaAs. (Reprinted from N. N. Ledentsov et al., *J. Cryst. Growth*, 301–302, 914 [2007]. With permission. Copyright 2007, Elsevier.)

8.6 Epitaxial Lateral Overgrowth

ELO,* now an important approach for mismatched heteroepitaxy, was originally developed for the fabrication of high-performance homoepitaxial devices in Si,[34] GaAs,[35,36] and InP.[37] In these applications of ELO, growth proceeds from seed windows cut through a mask layer (usually an oxide such as SiO_2). Its successful implementation requires selective growth (growth conditions that prevent nucleation and a deposition directly on the oxide). Then the growth over the oxide occurs entirely as an extension of the seed regions, resulting in a single-crystal layer. It should also be noted that the achievement of a planar layer requires a lateral growth rate that is much greater than the vertical rate (preferential growth). In principle, ELO can be applied to a number of heteroepitaxial material systems as long as the requirements of selective and preferential lateral growth can be met.

For the ELO of Si-on-Si(001) substrates with patterned SiO_2, selective epitaxial growth (SEG) is achieved by injecting HCl gas during growth from dichlorosilane ($SiCl_2H_2$).[38] Unfortunately, this process has a unity lateral-to-vertical growth rate ratio, resulting in a nonplanar surface.[39] (The maximum thickness grows over the seed windows, and the minimum thickness grows midway between seed windows.) This necessitates the use of chemical-mechanical polishing (CMP) to planarize the ELO material prior to device fabrication.[40]

The ELO growth of GaAs was first demonstrated by McClelland et al.,[35] using GaAs(110) substrates with a carbonized photoresist seed mask. The $AsCl_3$-GaAs-H_2 growth process (chloride VPE) was utilized, and the lateral-to-vertical growth rate ratio was approximately 25. This allowed the growth of GaAs layers having uniform thicknesses of 5–10 μm, which could be cleaved from the substrate, thus allowing its reuse. (This technique was termed the cleavage of lateral epitaxial films for transfer [CLEFT] process.[35]) Figure 8.16 shows this process in schematic fashion. Here, the use of chloride VPE gives preferential growth due to the difference in growth rates for different low-index faces, because the growth is kinetically controlled. It is therefore somewhat inflexible with regard to the choice of substrate orientation.

Gale et al.[36] demonstrated the ELO of GaAs using a SiO_2 mask and MOVPE. Selective growth was achieved without the use of HCl. The growth rate was preferential as well, with a lateral-to-vertical growth rate ratio of up to 5 on (110) substrates, achieved when the seed openings were misaligned from a $[\bar{1}10]$ direction by 2°–26°. Here, the growth conditions are such that the growth rate is minimum on low-index faces (facetted growth). Enhanced lateral growth occurs when the sidewalls are misoriented from a low-index crystal face. This can be understood as the consequence of a nearly unity sticking coefficient for the Ga precursor. Surface diffusion of the adsorbed Ga species then leads to enhanced growth on faces with high densities of steps and kinks.

Vohl et al.[37] studied the ELO of InP using the PCl_3-InP-H_2 (chloride VPE) process on InP substrates of different orientations. The growth was selective using a phosphosilicate glass (PSG) mask. Facetted growth resulted in the preferential growth at high-index faces.

Nam et al.[41] reported the first application of ELO for the attainment of continuous layers of GaN on mismatched heteroepitaxial substrates. In the selective growth of GaN hexagonal pyramids for field emitters on 6H-SiC(0001) substrates, they had discovered that unintended lateral growth occurred over the SiO_2 mask layers with certain growth conditions.[42] Moreover,

* This approach to heteroepitaxy also goes by the names *lateral epitaxial overgrowth* (LEO) and *selective area lateral epitaxial overgrowth* (SALEO).

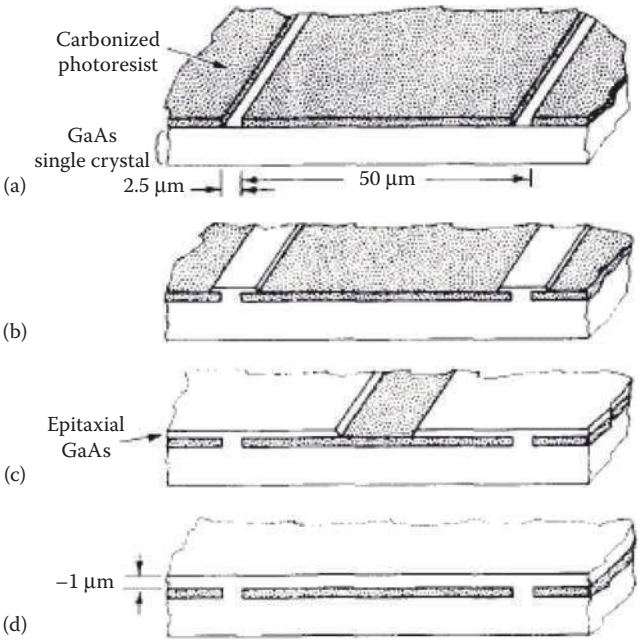


FIGURE 8.16
ELO of GaAs on a GaAs(110) substrate with a carbonized photoresist mask. (a) The mask is patterned with 2.5 μm wide slots, spaced 50 μm apart. (b, c) GaAs grows selectively in the slots (seed windows), and then grows laterally over the mask, with a lateral-to-vertical growth rate ratio of 25. (d) Adjacent areas of lateral growth merge to form a continuous layer of GaAs. (Reprinted from R. W. McClelland et al., *Appl. Phys. Lett.*, 37, 560 [1980]. With permission. Copyright 1980, American Institute of Physics.)

they found that the overgrown material contained a greatly reduced density of threading dislocations.[43] The reduction in the dislocation density in laterally overgrown GaN is shown dramatically in Figure 8.17. Here, GaN was grown laterally by MOVPE from a stripe-geometry seed region of GaN using a SiO_2 mask. The seed GaN was grown on a 6H-SiC(0001) substrate with a 1000 Å AlN buffer. Above the seed region, the threading dislocation density is 10^8 to 10^9 cm^{-2}, but there are no visible dislocations in the laterally grown material.

Nam et al.[41] then carried out a detailed investigation of ELO growth of GaN on vicinal 6H-SiC(0001) substrates, misoriented by 3°–4° toward the $\langle 11\bar{2}0\rangle$. In this study, the oxide mask openings were stripes aligned with the $\langle 1\bar{1}00\rangle$ and $\langle 11\bar{2}0\rangle$ directions. First, a 0.1 μm AlN buffer was grown on the 6H-SiC(0001) substrate by MOVPE (TEA + NH_3) at 1100°C, followed by a 1.5–2.0 μm thick layer of GaN grown by MOVPE (TEG + NH_3) at 1000°C. The 0.1 μm thick SiO_2 mask layer was deposited by low-pressure chemical vapor deposition at 410°C and patterned using photolithography and wet chemical etching in buffered HF. The stripe openings in the SiO_2 were oriented along the $\langle 1\bar{1}00\rangle$ and $\langle 11\bar{2}0\rangle$ directions, and were either 3 or 5 μm wide. The parallel stripes were spaced by distances of 3–40 μm. Following the patterning of the SiO_2 mask, and a dip in 50% buffered HCl to remove oxide from the exposed GaN surface, the lateral overgrowth of GaN was carried out by MOVPE (TEG + NH_3) at 1000°C–1100°C.

The morphology of the ELO GaN was very different for the $\langle 11\bar{2}0\rangle$ and $\langle 1\bar{1}00\rangle$ stripe orientations. Figure 8.18 shows scanning electron microscopy (SEM) micrographs of GaN grown on 3 μm wide stripe openings oriented along these two directions with various growth times. After only 3 min of growth, the morphology looks similar for the two stripe

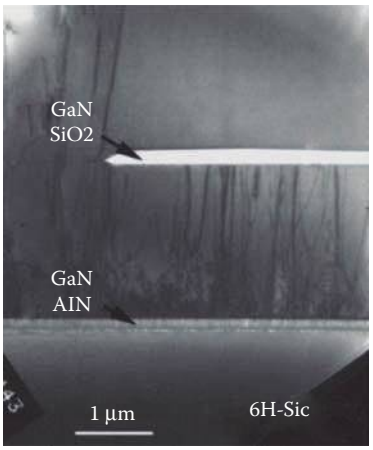


FIGURE 8.17
TEM micrograph in $[11\bar{2}0]$ orientation showing the reduction of the dislocation density in laterally grown GaN over a SiO_2 mask. The GaN was grown laterally by MOVPE from a stripe-geometry seed region of GaN using a SiO_2 mask. The seed GaN was grown on a 6H-SiC(0001) substrate with a 1000 Å AlN buffer. Above the seed region, the threading dislocation density is 10^8 to 10^9 cm^{-2}, but there are no visible dislocations in the laterally grown material. (Reprinted from T. S. Zheleva et al., *Appl. Phys. Lett.*, 71, 2472 [1997]. With permission. Copyright 1997, American Institute of Physics.)

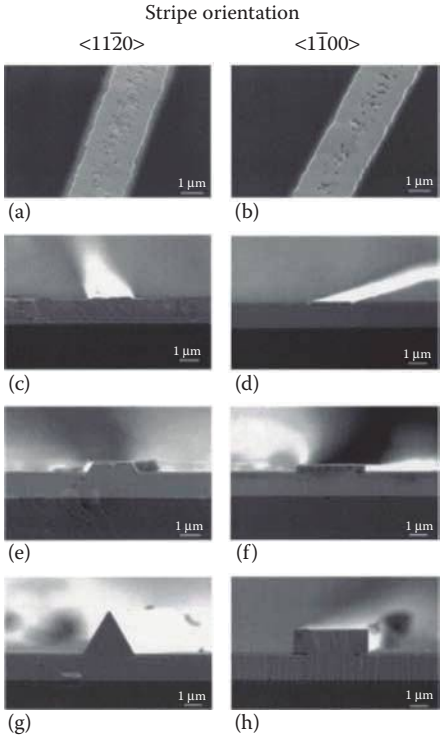


FIGURE 8.18
SEM micrographs of GaN grown on 3 μm wide oxide stripe openings oriented along the $\langle 11\bar{2}0\rangle$ and $\langle 1\bar{1}00\rangle$ directions with the growth times of (a–d) 3 min, (e, f) 9 min, and (g, h) 20 min. (Reprinted from O.-H. Nam et al., *Appl. Phys. Lett.*, 71, 2638 [1997]. With permission. Copyright 1997, American Institute of Physics.)

orientations. With additional growth, however, the stripes oriented along the $\langle 11\bar{2}0\rangle$ developed a triangular cross section with inclined $\{1\bar{1}01\}$ side facets. The stripes oriented along the $\langle 1\bar{1}00\rangle$, on the other hand, maintained a rectangular cross section with a (0001) top and {1120} sides.

Park et al.[44] further studied the effect of stripe orientation, using 3 μm wide stripe openings, 860 μm long, and indexed at 2° increments in a wagon wheel pattern, for the case of ELO GaN-on-sapphire (0001) substrates grown by MOVPE. They found the same cross sections and facets as Nam et al. for the $\langle 11\bar{2}0\rangle$ and $\langle 1\bar{1}00\rangle$ stripe orientations.

The lateral-to-vertical growth rate ratio for ELO GaN is also quite different for the $\langle 11\bar{2}0\rangle$ and $\langle 1\bar{1}00\rangle$ stripe orientations. Nam et al.[41] obtained a ratio less than unity for the $\langle 11\bar{2}0\rangle$-oriented stripes, and approximately unity for the $\langle 1\bar{1}00\rangle$ stripes. Park et al.[44] obtained a lateral-to-vertical growth rate ratio of up to 2 for $\langle 11\bar{2}0\rangle$-oriented stripes, as shown in Figure 8.19. They also found that the lateral-to-vertical growth rate ratio depends on the ratio of the "open" to "masked" stripe width (the fill factor) and the growth conditions, as well as the stripe orientation. As seen in Figure 8.19, the lateral-to-vertical growth ratio increases monotonically with the fill factor, for the range investigated.

Despite the relatively low lateral-to-vertical growth rate ratio, Nam et al. obtained smooth complete layers of ELO GaN by using stripes oriented along the $\langle 1\bar{1}00\rangle$. Figure 8.20 shows SEM micrographs of the cross section and top view for one such complete layer of ELO GaN, ~5 μm thick, grown using 3 μm stripes spaced by 3 μm. The surface of the coalesced ELO layer is relatively smooth. The 0.25 cm root mean square roughness is comparable with that for the underlying GaN layer. However, the process of coalescence leaves small voids above the oxide stripes, as can be seen in Figure 8.20.

The complete layers of GaN formed by ELO in this manner show a dramatic reduction in their threading dislocation density, in the laterally grown regions, as shown in the SEM micrograph of Figure 8.21 obtained by Nam et al. Above the seed stripe, threading

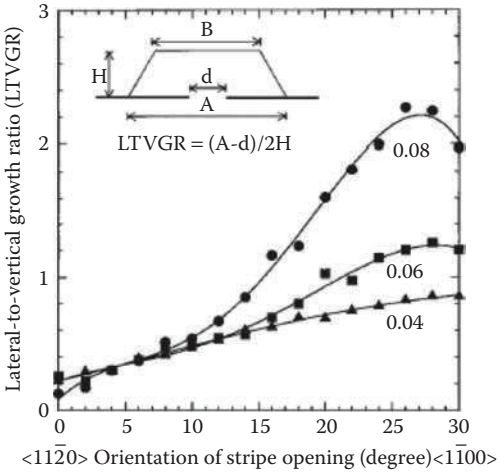
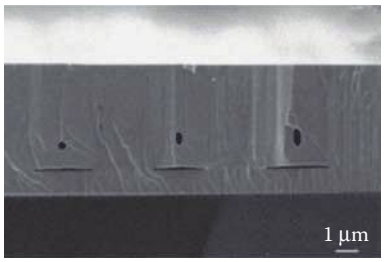
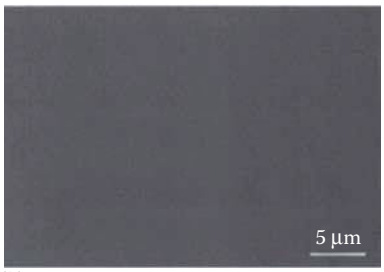


FIGURE 8.19
Lateral-to-vertical growth rate ratio for ELO GaN/AlN/α-Al_2O_3 (0001) grown by MOVPE, as a function of the stripe orientation with fill factor as a parameter. The fill factor is the ratio of open to masked surface area. The masking layer is 0.1 μm SiO_2. (Reprinted from J. Park et al., *Appl. Phys. Lett.*, 73, 333 [1998]. With permission. Copyright 1998, American Institute of Physics.)



(a)



(b)

FIGURE 8.20
SEM micrographs of the (a) cross section and (b) top view of a 5 μm complete layer of ELO GaN grown on vicinal 6H-SiC(0001) using an AlN buffer layer. The layer was grown using 3 μm oxide openings spaced by 3 μm, and oriented along the $\langle 1\bar{1}00 \rangle$ direction. (Reprinted from O.-H. Nam et al., *Appl. Phys. Lett.*, 71, 2638 [1997]. With permission. Copyright 1997, American Institute of Physics.)

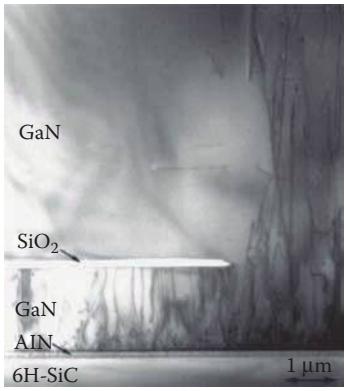


FIGURE 8.21
Cross-sectional TEM micrograph of ELO GaN grown on vicinal 6H-SiC(0001) using an AlN buffer layer. The layer was grown using 3 μm oxide openings spaced by 3 μm, and oriented along the $\langle 1\bar{1}00 \rangle$ direction. (Reprinted from O.-H. Nam et al., *Appl. Phys. Lett.*, 71, 2638 [1997]. With permission. Copyright 1997, American Institute of Physics.)

dislocations thread from the GaN/AlN interface to the surface of the GaN layer. In the laterally grown material, above the oxide mask, only dislocations parallel to the (0001) plane were visible. These are thought to be due to the bending over of threading dislocations from the highly defected region. However, they do not thread to the top surface. Moreover, in an investigation of Si-doped ELO GaN on sapphire, Park et al. found that the laterally

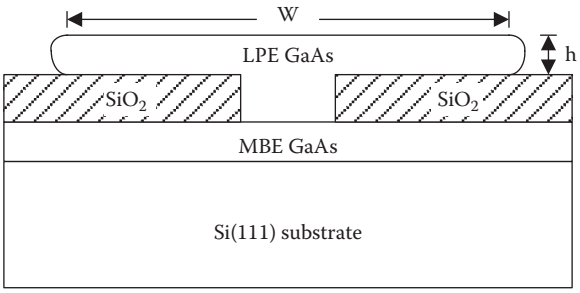


FIGURE 8.22
ELO (MCE) of GaAs on Si(111). First, a thin layer of GaAs is grown on the Si(111) by MBE. Then, an SiO_2 film is spun on and baked. A pattern of parallel line openings (microchannels) is produced in the SiO_2 by a photolithographic step. Finally, the ELO growth of GaAs is accomplished by LPE. A figure of merit for the process is the ratio of lateral-to-vertical growth. In their work, Chang et al. used the ratio W/h as defined in the figure.

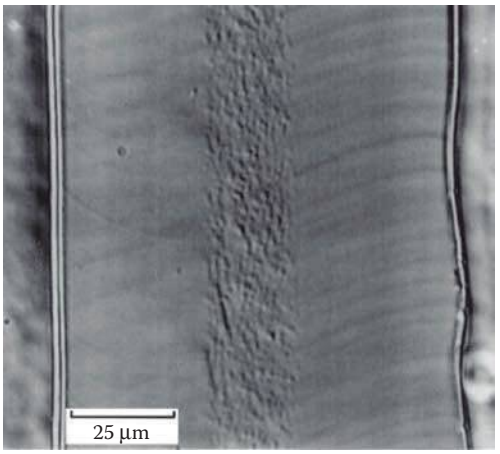


FIGURE 8.23
Etch pit morphology of GaAs on Si grown by ELO (MCE). The sample has been etched in molten KOH to reveal etch pits associated with threading dislocations. In the central region, where the GaAs grew vertically over the Si substrate, there is a high EPD. But the laterally grown regions, 47 μm on either side, are virtually free from etch pits. (Reprinted from Y. S. Chang et al., *J. Cryst. Growth*, 192, 18 [1998]. With permission. Copyright 1998, Elsevier.)

grown material exhibited two to three times the CL intensity of the material over the seed stripe (370 nm emission).

Chang et al.[45] investigated the ELO of GaAs-on-Si(111) substrates. In their approach, which they called "microchannel epitaxy" (MCE), they first grew a thin layer of GaAs on the Si(111) by MBE. Then a SiO_2 film was spun on and baked. A pattern of parallel line openings ("microchannels") was produced in the SiO_2 by a photolithographic step. The channel openings were 5 μm wide, and were separated by 200–1000 μm. Finally, the ELO growth of GaAs was accomplished by liquid phase epitaxy (LPE), resulting in the structure shown in Figure 8.22. Following ELO growth, the resulting material was characterized by crystallographic etching using molten KOH. In this study, it was found that the ratio of lateral-to-vertical growth could be as high as 17. As in the case of GaN, the ELO growth affected a dramatic reduction in the threading dislocation density in the laterally grown regions. This is shown in the optical micrograph of Figure 8.23, which shows the

EPD morphology of an MCE structure. In the central region, where the GaAs grew vertically over the Si substrate, there is a high EPD. But the laterally grown regions, 47 μm on either side, are virtually free from etch pits.

8.7 Pendeo-Epitaxy

Zheleva et al.[46] proposed "pendeo-epitaxy" as a new approach for the lateral growth of III-nitrides on mismatched heteroepitaxial substrates. This approach is similar to ELO, and makes use of the difference in growth rates for the (0001) and $\{11\bar{2}0\}$ planes. However, whereas ELO involves the lateral growth of GaN over a SiO_2 mask, from seed openings, the pendeo-epitaxy method involves lateral growth from mesa-patterned GaN and eliminates the need for a SiO_2 mask. Figure 8.24 shows both approaches schematically.

Zheleva et al. have discussed two modes of pendeo-epitaxial growth. In "mode A," the growth on the seed pillars proceeds faster in the lateral directions than in the vertical direction from its initiation. In "mode B," the growth is initially faster on the tops of the pillars, followed by rapid lateral growth from the newly formed $\{11\bar{2}0\}$ side facets. Figure 8.25 shows SEM and TEM micrographs of pendeo-epitaxial GaN grown in these two modes.

The primary advantage of pendeo-epitaxy compared with ELO is the elimination of the thermal strain associated with the SiO_2 mask. Like ELO, pendeo-epitaxy enables a dramatic decrease in the threading dislocation density (four to five orders of magnitude) in the laterally grown material, compared with direct growth. However, it has been found that the laterally grown wings exhibit crystallographic tilts with respect to the seed stripes, which are associated with the introduction of dislocations to relax the seed material.[47] This could lead to low-angle grain boundaries corresponding to the plane of wing coalescence.

8.8 Nanoheteroepitaxy

Nanoheteroepitaxy (NHE) is a substrate patterning approach that involves the growth of nanometer-scale islands on a mismatched heteroepitaxial substrate. In the implementation of NHE, selective epitaxial growth is carried out on a substrate that has been patterned to have nanometer-scale seed pads. This may be achieved either by etching windows through a dielectric mask material or by mesa etching the substrate crystal. Typically, lateral epitaxial growth proceeds from the seed pads until the coalescence of the growing islands yields a complete layer of the heteroepitaxial material. NHE differs from ELO in that the pattern involves islands rather than stripes, and that the seed pads have dimensions on the order of nanometers, not micrometers. This latter feature is not simply a difference in degree, but introduces new mechanisms of strain relaxation that are of fundamental importance. This is illustrated in Figure 8.26 for the case of a mesa-patterned substrate. In the planar growth of a pseudomorphic mismatched layer, as shown on the left, the only stress relief mechanism other than the creation of dislocations is the vertical deformation of the epitaxial material (a). However, for the mismatched material on the mesa-patterned substrate, the stress can also be relieved by lateral deformations in the epitaxial layer (b), along with vertical and lateral deformations in the substrate mesas (c and d). The stress

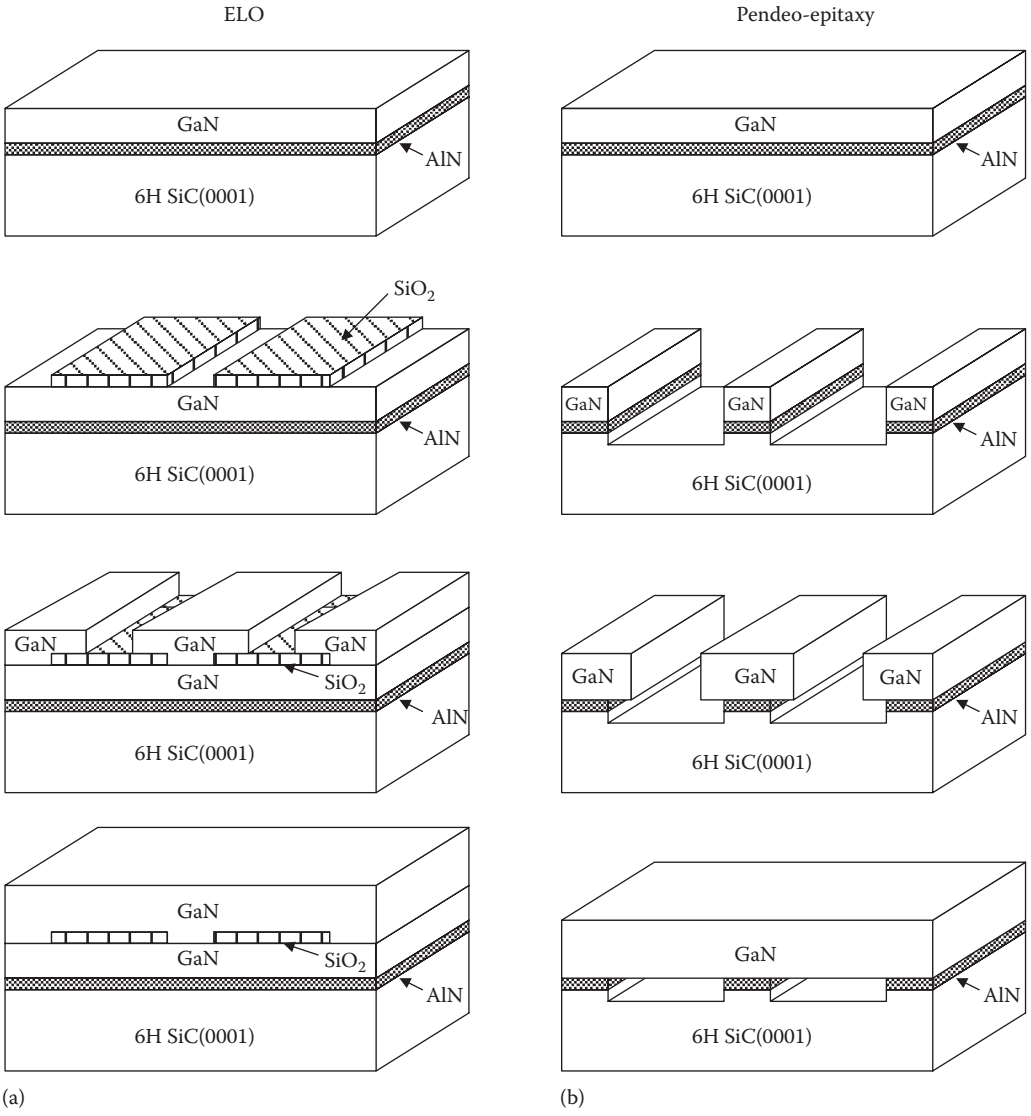


FIGURE 8.24
ELO and pendeo-epitaxy approaches for the growth of high-quality GaN-on-SiC or sapphire substrates. (a) ELO process: An AlN buffer is grown. A GaN layer is grown. SiO_2 is deposited and patterned to open seed stripes. GaN is grown from the seed stripes laterally over the remaining SiO_2 to produce a complete layer. (b) Pendeo-epitaxy process: An AlN buffer is grown. A GaN layer is grown. The GaN/AlN is mesa dry etched to create seed pillars. GaN is grown laterally from the pillars, to create a complete layer.

relief mechanisms (b–d) that are unique to patterned growth can make it possible to grow the heteroepitaxial material much thicker than the Matthews and Blakeslee critical layer thickness for planar growth, without the introduction of misfit dislocations. Even if stress relief should occur partly by the introduction of misfit dislocations, the associated threading dislocations can easily glide to the edge of the mesa, as in the PHeP approach. Whereas the planar layer will contain threading dislocations such as (e), which are unable to glide to

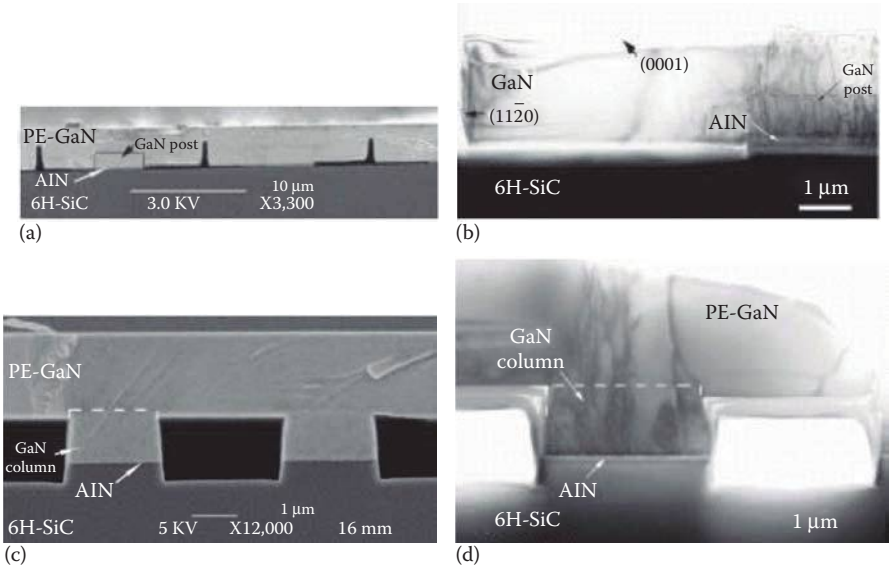


FIGURE 8.25
(a) SEM and (b) TEM micrographs of pendeo-epitaxial GaN grown in mode A; (c) SEM and (d) TEM micrographs of pendeo-epitaxial GaN grown in mode B. (Reprinted from T. Zheleva et al., *MRS Internet J. Nitride Semicond.*, 4S1, G3.38 [1999]. With permission. Copyright 1999, Materials Research Society.)

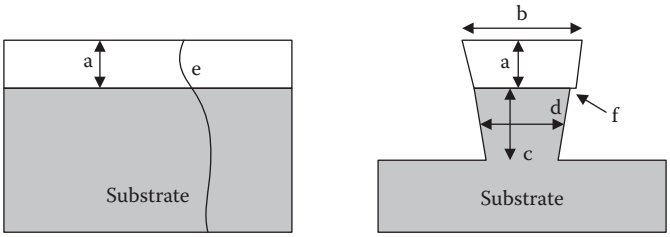


FIGURE 8.26
Stress relief mechanisms in a mismatched heteroepitaxial crystal. For a planar layer as shown on the left, the only stress relief mechanism is the vertical deformation of the epitaxial material (a). But in the case of NHE, shown on the right, the epitaxial deposit can also deform in the lateral direction (b). In addition, the substrate mesas can deform both vertically (c) and laterally (d). For planar heteroepitaxy, a substrate dislocation (e) will thread through the epitaxial layer, but in the case of NHE, the threading dislocation can glide to a sidewall to create a sidewall step (f), similar to the case of PHeP.

the edge of the sample, the NHE layer is likely to contain only misfit dislocations and their associated sidewall steps, as in (f).

Luryi and Suhir[48] first presented a theoretical treatment of strain relaxation in a heteroepitaxial layer on a nanopatterned substrate. They considered growth of a mismatched semiconductor that makes rigid contact with the substrate only on round seed pads, the diameter of which is on the scale of nanometers. They showed that, in a pseudomorphic structure of this sort, the strain in the heteroepitaxial layer decays exponentially with distance from the interface. The characteristic length h_e for this decay is on the order of the seed pad diameter. Because of this behavior, the critical layer thickness increases as the seed pads are scaled down in size. For a given lattice mismatch strain, there is a seed pad

size below which the critical layer thickness diverges to infinity, so that structures entirely free from misfit dislocations may be produced.

In Section 8.8.1, the Luryi and Suhir model is outlined for the case of a noncompliant substrate. The section describes the extension of this theory to include substrate compliancy, as developed by Zubia and Hersee. Finally, experimental results for NHE are summarized.

8.8.1 Nanoheteroepitaxy on a Noncompliant Substrate

Luryi and Suhir[48] developed the first theoretical model for the strain in nanoheteroepitaxial material. They assumed that the lattice-mismatched heteroepitaxial material makes rigid contact with a noncompliant substrate only at round seed pads having a diameter 2*l*, as shown in Figure 8.27. Here, the *y* axis lies in the plane of the interface, along a major cord of a seed pad. The *z* axis is perpendicular to the substrate and passes through the center of this seed pad. Figure 8.27 shows a heteroepitaxial layer that has coalesced into a single layer by lateral growth, and the total thickness of the heteroepitaxial layer is *h*. It is assumed that the areas between the seed pads are wide enough so there is no interference of the strain fields from adjacent pads. In this situation, if the substrate is unstrained, then the in-plane stress in the epitaxial deposit is given by

$$\sigma_{||} = f\frac{E}{1-\nu}\chi(y,z)\exp(-\pi z/2l), \tag{8.8}$$

where:

- f is the lattice mismatch strain
- E is the Young's modulus
- ν is the Poisson ratio

$$\chi(y,z) = \begin{cases} 1-\dfrac{\cosh(ky)}{\cosh(kl)}; & z \le h_e \\ 1; & z \ge h_e \end{cases}, \tag{8.9}$$

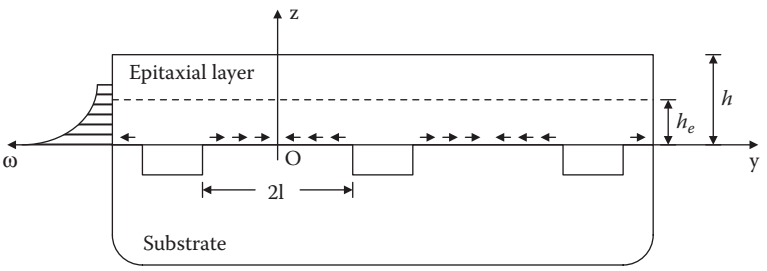


FIGURE 8.27
Nanoheteroepitaxial growth on a patterned substrate. The substrate has been patterned with round seed pads having a diameter 2*l*. The *y* axis lies in the plane of the interface, along a major cord of one of the seed pads. The *z* axis is perpendicular to the substrate and passes through the center of this seed pad. The heteroepitaxial layer may coalesce into a single layer by lateral growth, as shown. The total thickness of the heteroepitaxial layer is *h*. (Adapted from S. Luryi and E. Suhir, *Appl. Phys. Lett.*, 49, 140 [1986]. With permission. Copyright 1986, American Institute of Physics.)

where h_e is the effective range for the stress in the z direction, to be determined below, and the interfacial compliance parameter k is given by

$$k = \left[\frac{3}{2}\left(\frac{1-\nu}{1+\nu}\right)\right]^{1/2}\frac{1}{h_e} \equiv \frac{\zeta}{h_e}\,. \tag{8.10}$$

The strain energy density per unit volume is

$$\omega(y,z) = \frac{1-\nu}{E}\sigma_{||}{}^2\,, \tag{8.11}$$

and is maximum at $y=0$. The strain energy per unit area may be found by integrating over the thickness of the epitaxial deposit and takes on a maximum value at $y=0$, which is

$$E_e = \int_0^h \omega(0,z) \equiv \frac{E}{1-\nu} f^2 h_e{}^2\,. \tag{8.12}$$

In this calculation, there is little contribution from $z > h_e$, so that it is a good approximation to use the form of $\chi(y, z)$ for $z \le h_e$. The right-hand side of Equation 8.12 defines the characteristic thickness h_e, which is then given implicitly by

$$h_e = h\left\{\left[1-\sec h\left(\frac{\zeta l}{h_e}\right)\right]^2\left[1-\exp\left(-\pi h/l\right)\right]\frac{l}{\pi h}\right\} = h\left[\phi\left(\frac{l}{h}\right)\right]^2\,, \tag{8.13}$$

where

$$\zeta = \left[\frac{3}{2}\left(\frac{1-\nu}{1+\nu}\right)\right]^{1/2}\,. \tag{8.14}$$

Numerical solution of Equation 8.14 yields the reduction factor $\phi(l/h)$, which is plotted in Figure 8.28. For $l \gg h$, $\phi \to 1$ asymptotically, but for $l \ll h$, $\phi \propto (l/h)^{1/2}$. In other words, $h_e \approx h$ for $l \gg$ h, and for $l \ll$ h,

$$h_e \approx h\left[1-\sec h(\zeta\pi)\right]^2\,. \tag{8.15}$$

The strain energy per unit area from Equation 8.12 may be used in conjunction with an energy calculation for the critical layer thickness, to find the critical layer thickness h_c^l for an island of radius l. The result is

$$h_c^l = h_c\left[\phi\left(l/h_c^l\right)f\right]\,.$$

In their work, Luryi and Suhir used the People and Bean model for the determination of the critical layer thickness, $h_c[x] = 0.1\left(a_0/|x|\right)$. However, the Matthews energy calculation of the critical layer thickness may also be used, with

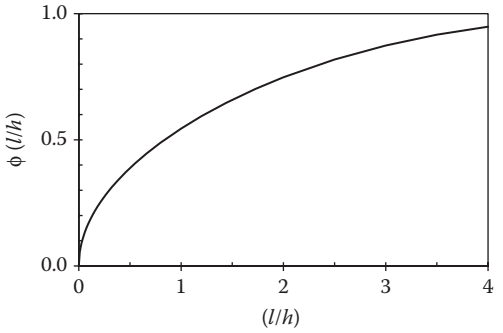


FIGURE 8.28
Reduction factor ϕ as a function of l/h, where l is the radius of the seed pads and h is the epitaxial layer thickness. (Adapted from S. Luryi and E. Suhir, *Appl. Phys. Lett.*, 49, 140 [1986]. With permission. Copyright 1986, American Institute of Physics.)

$$h_c(x) = \frac{b\left(1-\nu\cos^2\alpha\right)\left[\ln\left(h_c(x)/b\right)+1\right]}{8\pi|f|(1+\nu)\cos\lambda}. \tag{8.16}$$

8.8.2 Nanoheteroepitaxy with a Compliant Substrate

Zubia and Hersee[49] extended this theory to include the effect of substrate compliance and named the approach "nanoheteroepitaxy." Here, the strain is partitioned between the substrate and the epitaxial layer. If the epitaxial layer is grown coherently (without misfit dislocations) on a compliant substrate with lattice mismatch strain f, then the substrate and epitaxial layer will be strained in opposite sense such that

$$\varepsilon_{epi} - \varepsilon_{sub} = f, \tag{8.17}$$

where ε_{epi} and ε_{sub} are the in-plane strains in the epitaxial layer and the substrate, respectively. If we neglect the bending stresses, force balance in the structure dictates that[50]

$$\sigma_{epi}h_{epi} + \sigma_{sub}h_{sub} = 0, \tag{8.18}$$

where:

h_{epi} and h_{sub} are the thicknesses of the epitaxial layer and substrate
σ_{epi} and σ_{sub} are the corresponding in-plane stresses

Due to the biaxial nature of the stress, the stress–strain relationships are

$$\sigma_{epi} = \frac{E_{epi}}{1-\nu_{epi}}\varepsilon_{epi}$$

and

$$\sigma_{sub} = \frac{E_{sub}}{1-\nu_{sub}}\varepsilon_{sub}, \tag{8.19}$$

where:
E_{epi} and E_{sub} are the Young's moduli
ν_{epi} and ν_{sub} are the Poisson ratios

The simultaneous solution of these three equations yields

$$\varepsilon_{epi} = \frac{f}{1+\left(K\dfrac{h_{epi}}{h_{sub}}\right)}$$

and

$$\varepsilon_{sub} = \frac{-f}{1+\left(\dfrac{1}{K}\dfrac{h_{sub}}{h_{epi}}\right)}, \tag{8.20}$$

where k is given by

$$K = \frac{E_{epi}}{\left(1-\nu_{epi}\right)}\frac{\left(1-\nu_{sub}\right)}{E_{sub}}. \tag{8.21}$$

Now, combining the compliant substrate theory with the model of Luryi and Suhir, we have

$$\sigma_{epi} = \varepsilon_{epi}\frac{E_{epi}}{1-\nu_{epi}}\chi(y,z)\exp\left(-\frac{\pi z}{2l}\right) \tag{8.22}$$

and

$$\sigma_{sub} = \varepsilon_{sub}\frac{E_{sub}}{1-\nu_{sub}}\chi(y,z)\exp\left(-\frac{\pi z}{2l}\right). \tag{8.23}$$

The in-plane strains in the nanoheteroepitaxial material and substrate pads are given by

$$\varepsilon_{epi} = \frac{f}{1+\left(K\dfrac{\left(1-\exp\left(-\pi h_{epi}/2l\right)\right)}{\left(1-\exp\left(-\pi h_{sub}/2l\right)\right)}\right)} \tag{8.24}$$

and

$$\varepsilon_{sub} = \frac{-f}{1+\left(\frac{1}{K}\frac{\left(1-\exp\left(-\pi h_{sub}/2l\right)\right)}{\left(1-\exp\left(-\pi h_{epi}/2l\right)\right)}\right)}. \tag{8.25}$$

Figure 8.29 shows the partitioning of strain between an epitaxial layer and compliant substrate, for the case of planar (unpatterned) growth. It can be seen that more of the strain is transferred to the substrate if the thickness ratio h_{epi}/h_{sub} is increased or if K is increased.

When the strain partitioning is accounted for, along with the stress relief by lateral contraction or expansion at the sidewalls, it is expected that layers can be grown coherently much greater than the critical layer thickness for one-dimensional (planar) growth (as shown Figure 8.30). In fact, Zubia and Hersee predicted it should be possible to grow layers with a lattice mismatch of 4.2% completely dislocation-free.

Zubia et al.[51] investigated the NHE of GaN on patterned silicon-on-insulator (SOI) by MOVPE. The SOI (111) wafers were produced using the separation by ion implantation of oxygen (SIMOX) process. The SOI wafers were patterned using interferometric photolithography and reactive ion etching in the manner reported by Zaidi et al.,[52,53] forming a square two-dimensional array of silicon islands on top of SiO_2. These silicon islands had a height of 40 nm and diameters of 80–300 nm, and their separation in the square array was 360 or 900 nm. An interesting aspect of this work was the reduced melting point of the silicon islands. A comparison of their morphology before and after heating to 1110°C in the epitaxial reactor (Figure 8.31) shows that the nanoscale silicon islands melted at or below this temperature, even though the bulk melting temperature is 1412°C.

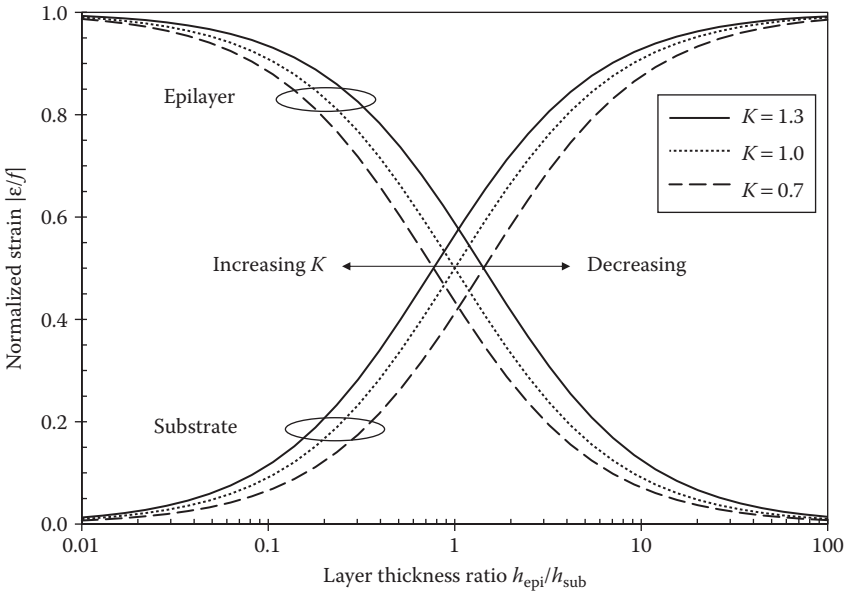


FIGURE 8.29
Strain partitioning between an epitaxial layer and compliant substrate, for the case of planar (unpatterned) growth. (Adapted from D. Zubia and S. D. Hersee, *J. Appl. Phys.*, 85, 6492 [1999]. With permission. Copyright 1999, American Institute of Physics.)

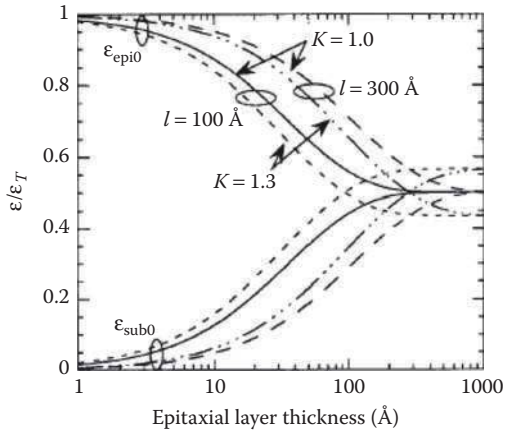


FIGURE 8.30
Strain partitioning for NHE. The nanoisland diameter is $2l$ and the substrate thickness was assumed to be 50 nm. (Reprinted from D. Zubia and S. D. Hersee, *J. Appl. Phys.*, 85, 6492 [1999]. With permission. Copyright 1999, American Institute of Physics.)

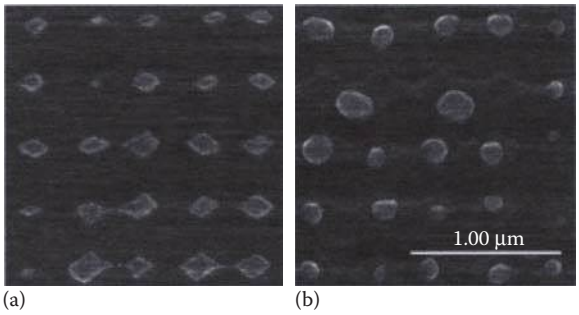


FIGURE 8.31
Morphology of silicon nanoscale islands (a) before and (b) after heating to 1110°C in the epitaxial reactor. The change in shape shows that the nanoscale islands had melted at or below this temperature. (Reprinted from D. Zubia et al., *Appl. Phys. Lett.*, 76, 858 [2000]. With permission. Copyright 2000, American Institute of Physics.)

The melting point reduction of the nanoscale islands can give rise to their softening, and an enhanced compliant substrate effect for NHE on top of them. Zubia et al. referred to this effect as *active compliance.* This effect was not included in the original model for NHE, but it can be accounted for by the use of an effective value of K.

Figure 8.32 shows cross-sectional TEM micrographs of nanoheteroepitaxial GaN on SOI, for the case of 80 and 280 nm islands. In both cases, the nanoheteroepitaxial GaN contained dislocations near the interface. However, the dislocation density decreased with distance from the interface.

8.9 Planar Compliant Substrates

In planar-mismatched heteroepitaxy on a thick substrate, all of the mismatch strain resides in the epitaxial layer, which must be less than the critical layer thickness to avoid

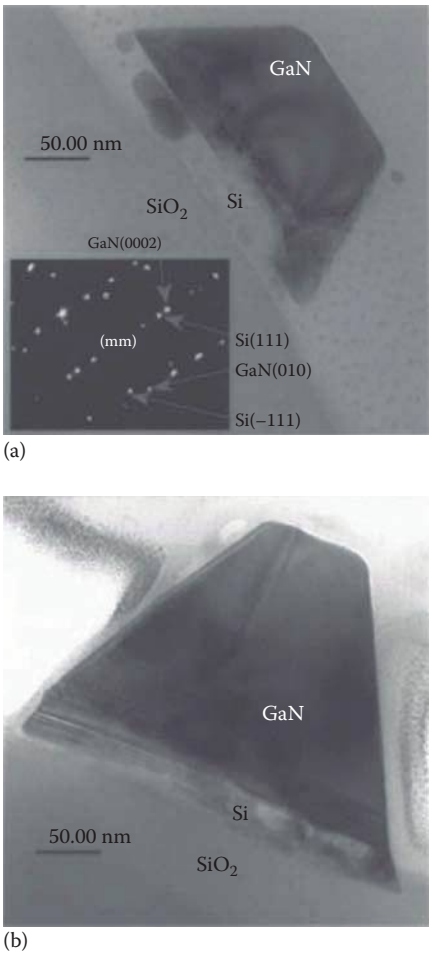


FIGURE 8.32
Cross-sectional TEM micrographs of nanoheteroepitaxial GaN grown on SOI islands with diameters of (a) 80 nm and (b) 280 nm. (Reprinted from D. Zubia et al., *Appl. Phys. Lett.*, 76, 858 [2000]. With permission. Copyright 2000, American Institute of Physics.)

the introduction of dislocations. However, pseudomorphic layers thicker than the critical layer thickness would be beneficial in many device applications. In order to lift the critical layer thickness constraint for pseudomorphic growth, Lo[54] proposed the use of compliant substrates.

A compliant substrate is one thin enough so that it becomes strained by the deposition of a mismatched heteroepitaxial layer.* The partitioning of strain between the epitaxial layer and substrate reduces the total strain energy. If the substrate is sufficiently thin, the overall strain energy will never be large enough to cause the production of misfit dislocations. Then, the effective critical layer thickness will diverge to infinity so that a pseudomorphic layer of any thickness may be grown.

* In practical implementations of compliant substrates, the thin template layer is not free standing, but mechanically decoupled from a thick handle wafer.

In practice, a membrane thin enough to act as a compliant substrate is difficult to handle during processing. Another problem is that compliant membranes are susceptible to bowing and other distortions during strained heteroepitaxy. For these reasons, the use of compliant substrate technology requires the realization of a thin compliant layer on a rigid handle wafer. In such a realization, the handle wafer must restrain the compliant layer in the growth direction, to prevent buckling. However, the compliant layer must be mechanically decoupled from the substrate in the plane of the interface. No perfect scheme for a large-area compliant substrate on a handle layer has been demonstrated. On the other hand, compliant substrate technologies of this general type have been investigated with various degrees of success. These approaches include glass-bonded, metal-bonded, and twist-bonded wafers, and SOI.

In Section 8.8, we considered briefly the theory of compliant substrates in the context of NHE. Here, we will consider compliant substrate theory in more detail, for its application to one-dimensional (unpatterned) heteroepitaxy. We will also review various schemes for compliant substrate realization, along with descriptions of the relevant experimental observations.

8.9.1 Compliant Substrate Theory

If an epitaxial layer is grown coherently (without misfit dislocations) on a compliant substrate with lattice mismatch strain f, then the substrate and epitaxial layer will be strained in opposite sense such that

$$\varepsilon_{epi} - \varepsilon_{sub} = f, \tag{8.26}$$

where ε_{epi} and ε_{sub} are the in-plane strains in the epitaxial layer and the substrate, respectively. If we neglect the bending stresses, force balance in the structure dictates that[50]

$$\sigma_{epi} h_{epi} + \sigma_{sub} h_{sub} = 0, \tag{8.27}$$

where:

h_{epi} and h_{sub} are the thicknesses of the epitaxial layer and substrate
σ_{epi} and σ_{sub} are the corresponding in-plane stresses

Due to the biaxial nature of the stress, the stress–strain relationships are

$$\sigma_{epi} = \frac{E_{epi}}{1 - \nu_{epi}} \varepsilon_{epi}, \tag{8.28}$$

and

$$\sigma_{sub} = \frac{E_{sub}}{1 - \nu_{sub}} \varepsilon_{sub}, \tag{8.29}$$

where:

E_{epi} and E_{sub} are the Young's moduli
ν_{epi} and ν_{sub} are the Poisson ratios

The simultaneous solution of these three equations yields

$$\varepsilon_{\text{epi}} = \frac{f}{1+\left(K\dfrac{h_{\text{epi}}}{h_{\text{sub}}}\right)}, \tag{8.30}$$

and

$$\varepsilon_{\text{sub}} = \frac{-f}{1+\left(\dfrac{1}{K}\dfrac{h_{\text{sub}}}{h_{\text{epi}}}\right)}, \tag{8.31}$$

where K is given by

$$K = \frac{E_{\text{epi}}}{\left(1-\nu_{\text{epi}}\right)}\frac{\left(1-\nu_{\text{sub}}\right)}{E_{\text{sub}}}. \tag{8.32}$$

The strain energy per unit area in the bilayer system is

$$E_e = \frac{E_{\text{epi}}}{\left(1-\nu_{\text{epi}}\right)}h_{\text{epi}}\varepsilon_{\text{epi}}{}^2 + \frac{E_{\text{sub}}}{\left(1-\nu_{\text{sub}}\right)}h_{\text{sub}}\varepsilon_{\text{sub}}{}^2. \tag{8.33}$$

With the approximations $E_{\text{epi}}=E_{\text{sub}}=E$ and $\nu_{\text{epi}}=\nu_{\text{sub}}=\nu$, the strain energy simplifies to

$$E_e = \frac{E}{\left(1-\nu\right)}f^2\left(\frac{h_{\text{epi}}h_{\text{sub}}}{h_{\text{epi}}+h_{\text{sub}}}\right). \tag{8.34}$$

If it is assumed that misfit dislocations will be introduced when the areal strain energy exceeds the misfit dislocation energy per unit area E_d as determined by Matthews, then the effective critical layer thickness h_{eff} for the epitaxial layer on the compliant substrate of thickness h_{sub} can be found from

$$\frac{1}{h_{\text{eff}}} = \frac{1}{h_c} - \frac{1}{h_{\text{sub}}}, \tag{8.35}$$

where h_c is the Matthews and Blakeslee critical layer thickness. Figure 8.33 shows the normalized critical layer thickness h_{eff}/h_c versus the normalized substrate thickness h_{sub}/h_c. It is important to note that when $h_{\text{sub}}<h_c$, there exists no solution to Equation 8.35 so that $h_{\text{eff}}\rightarrow\infty$.

Lo[54] also showed that, in partially relaxed epitaxial layers on compliant substrates, the modified image forces could help reduce the density of threading dislocations in the epitaxial layer. For the case of a thick substrate, the image force always attracts dislocations toward the free surface of the epitaxial layer. In that situation, the image force is associated with the free surface of the epitaxial layer, and it is equal to the attractive force that would

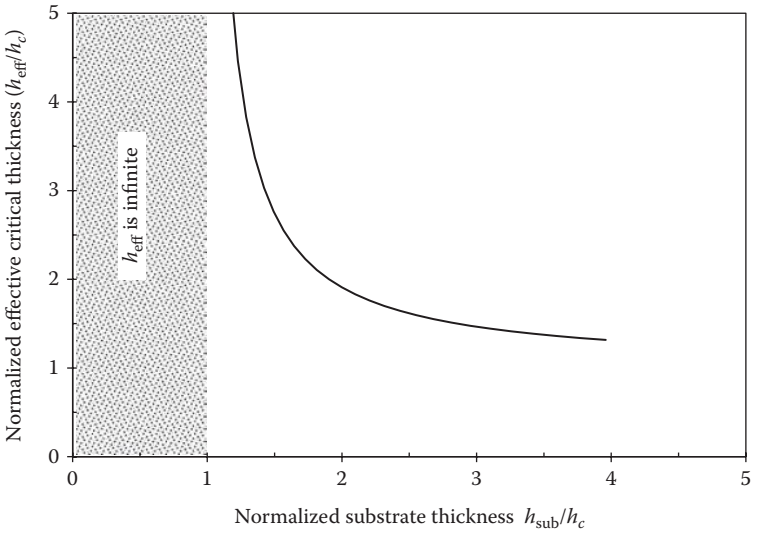


FIGURE 8.33
Normalized critical layer thickness h_{eff}/h_c vs the normalized substrate thickness h_{sub}/h_c for the growth of a mismatched heteroepitaxial layer on a compliant substrate. h_{eff} is the effective critical layer thickness, h_{sub} is the thickness of the compliant substrate, and h_c is the Matthews and Blakeslee critical layer thickness. (Adapted from Y. H. Lo, *Appl. Phys. Lett.*, 59, 2311 [2005]. With permission. Copyright 2005, American Institute of Physics.)

exist between the real dislocation and an image dislocation, with the opposite Burgers vector, and located at an equal distance from the surface but on the opposite side of it. With a compliant substrate, the image force can be greatly decreased in magnitude, or may even change sign and drive the dislocation into the substrate, away from the epitaxial layer surface. Here, both free surfaces contribute to the overall image force. Lo calculated this image force for a 60° misfit dislocation along a ⟨110⟩ direction for (001) heteroepitaxy of a zinc blende material on a compliant substrate. The image force per unit length of dislocation was found to be

$$\frac{F_I}{L} = \frac{Gb^2}{4}\left(\frac{1}{4} + \frac{1}{2(1-\nu)}\right)\left(\frac{1}{h_{epi} + h_{sub}}\right)\cot\left(\frac{\pi h_{epi}}{h_{epi} + h_{sub}}\right). \tag{8.36}$$

The misfit dislocations experience the maximum attractive image force in the case of an infinite (noncompliant) substrate. A reduction in the substrate thickness decreases the attractive image force, for a given thickness of the epitaxial layer. The sign of the image force may even change, indicating that the dislocations will be repelled from the surface of the epitaxial layer (or attracted to the surface of the compliant substrate). This behavior is illustrated in Figure 8.34, which shows the image force (arbitrary units) versus the normalized epitaxial layer thickness h_{epi}/h_c, with the normalized substrate thickness h_{sub}/h_c as a parameter.

8.9.2 Compliant Substrate Implementation

Several approaches have been invented for the implementation of compliant substrates. The first implementation of a compliant substrate was the cantilevered membrane, proposed

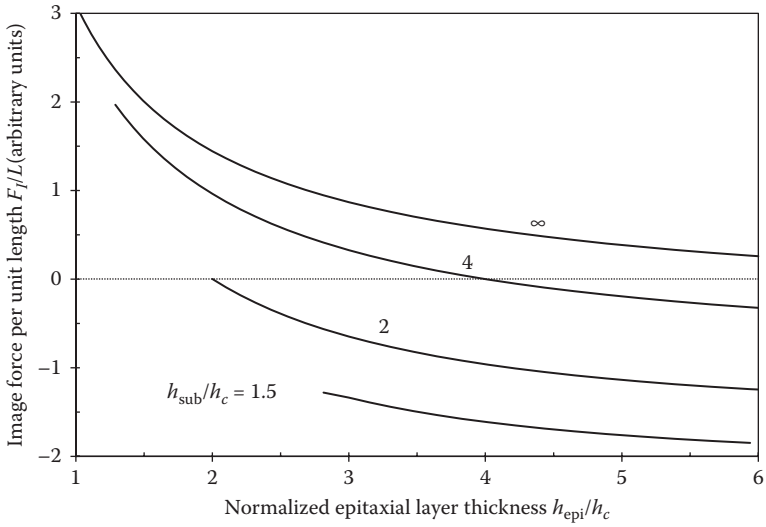


FIGURE 8.34
Image force (arbitrary units) for a 60° misfit dislocation at the interface between an epitaxial layer and a compliant substrate, vs. the normalized epitaxial layer thickness h_{epi}/h_c, and with the normalized substrate thickness h_{sub}/h_c as a parameter, where h_c is the Matthews and Blakeslee critical layer thickness. (Adapted from Y. H. Lo, *Appl. Phys. Lett.*, 59, 2311 [2005]. With permission. Copyright 2005, American Institute of Physics.)

by Teng and Lo[55] and demonstrated by Chua et al.[56] Other approaches have involved the realization of a thin compliant layer, which is on top of a thick handle wafer but mechanically decoupled from it. Along these lines, wafer bonding is the most studied method. Here, an etch stop layer and compliant layer are grown epitaxially on one wafer, which is then bonded to a handle wafer. The former wafer is then removed by lapping and etching, leaving just the compliant layer bonded to the handle wafer. Some degree of compliancy is achieved by the use of an intermediate layer (e.g., metal or glass) between the compliant layer and the handle wafer, or by a twist bond. Carter-Coman et al.[57,58] developed compliant substrates using wafer bonding with an intermediate layer of indium. This indium layer melts at epitaxial growth temperatures, rendering the thin layer compliant. Moran et al.[59,60] have developed bonded compliant substrates using intermediate layers of borosilicate glass. In the case of twist-bonded compliant substrates, there is no need for the insertion of an intermediate layer. Instead, the twist boundary introduces an array of screw dislocations, which introduces some level of compliancy. SOI wafers have also been investigated as potential compliant substrates. For example, Powell et al.[61] studied the epitaxy of SiGe alloys on an ultrathin SOI layer. Yang et al.[62] extended this work to the growth of GaN on both SOI substrates and SiC-on-SOI substrates. However, Rehder et al.[63] showed that, in the case of SiGe heteroepitaxy, a thin SOI layer does not act as a compliant substrate (in the sense of strain partitioning), even though it does alter the dislocation structure and dynamics.

In the following sections, the various approaches for realizing compliant substrates will be described in detail, including cantilevered membranes; glass-bonded, metal-bonded, and twist-bonded wafers; and SOI. Future directions will also be discussed, along with an introduction to the proposed universal compliant trench (UCT) substrate.

8.9.2.1 Cantilevered Membranes

Teng and Lo[55] proposed the use of a cantilevered membrane as a compliant substrate. Their design, shown in Figure 8.35, could be created by an undercutting etch. The membrane is supported at the four corners, but should behave as a compliant substrate in the central region away from the supports. They were able to fabricate such cantilevered membranes using selective wet etching of GaAs/AlGaAs and InP/InGaAs epitaxial structures.

Chua et al.[56] demonstrated the use of a cantilevered membrane having a bench structure as shown in Figure 8.36. To create the 80 nm membrane, they first grew a 80 nm GaAs/100 nm $Al_{0.8}Ga_{0.2}As$/GaAs(001) structure by MBE. Then, using photolithography and a nonselective etch, they mesa-etched stripes, 5 μm wide and with a 10 μm center-to-center spacing, in the 80 nm GaAs/100 nm $Al_{0.8}Ga_{0.2}As$. With a second photolithographic step, they opened stripes perpendicular to the first set, having a width of 10 μm, and etched the $Al_{0.8}Ga_{0.2}As$ selectively through these openings with 1 HF:5 H_2O.

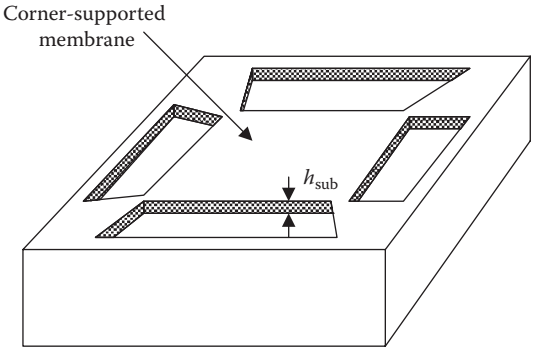


FIGURE 8.35
Cantilevered membrane for use as a compliant substrate. (Adapted from D. Teng and Y. H. Lo, *Appl. Phys. Lett.*, 62, 43 [1993]. With permission. Copyright 1993, American Institute of Physics.)

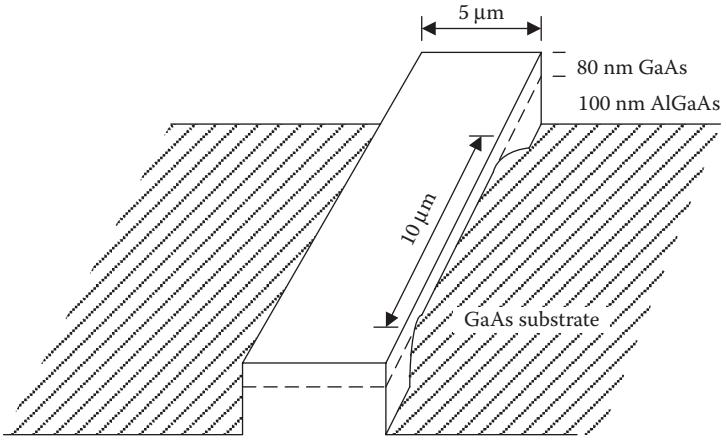


FIGURE 8.36
Cantilevered membrane with a bench structure, for use as a compliant substrate. (Adapted from C. L. Chua et al., *Appl. Phys. Lett.*, 64, 3640 [1994]. With permission. Copyright 1994, American Institute of Physics.)

On the cantilevered membrane described above, Chua et al. grew $In_{0.14}Ga_{0.86}As$, for which the room temperature mismatch strain is $f = 0.94\%$ and the critical layer thickness is $h_c \approx 10$ nm. They grew $In_{0.14}Ga_{0.86}As$, 200 cm thick, simultaneously on the compliant platform and on a reference, unprocessed GaAs substrate. The $In_{0.14}Ga_{0.86}As$ is about 20 times the expected critical layer thickness, as predicted by the Matthews and Blakeslee model. They found that the 004 XRD peak separation (between the XRD peaks for the GaAs and $In_{0.14}Ga_{0.86}As$) was significantly greater on the compliant platform than on the reference substrate. This could be interpreted as an indication of tetragonal distortion in the GaAs compliant platform, which would be expected if it is compliant. Chua et al. also studied the surface morphology of the $In_{0.14}Ga_{0.86}As$ by atomic force microscopy (AFM). The layer on the reference substrate showed a rough surface morphology associated with dislocations introduced during lattice relaxation. In contrast, the layer grown on the compliant platform exhibited a smooth surface texture.

Despite these results, cantilevered compliant membranes are expected to be mechanically fragile and will have poor heat removal performance. Other approaches, which produce a thin compliant layer on a handle wafer, are therefore preferred.

8.9.2.2 Silicon-on-Insulator as a Compliant Substrate

SOI has been investigated for use as a compliant substrate for the growth of SiGe alloys, GaN, and GaAs. In the case of an SOI wafer, the silicon layer may act as a compliant substrate if it is sufficiently thin, and mechanically decoupled from the wafer by slippage at the Si/SiO_2 interface. Experimental results with this type of compliant substrate have been mixed, however. Recent experiments by Rehder et al.[63] involving the growth of SiGe-on-SOI wafers indicate that the silicon layer does not behave as a compliant substrate in the usual sense. On the other hand, Rehder et al. and others have measured the existence of partial strain partitioning between the thin Si layer and the SiGe, and observed the preferential introduction of dislocations in the Si rather than the SiGe.

Powell et al.[61] performed initial experiments with the use of an ultrathin SOI layer as a compliant substrate, for the epitaxy of SiGe alloys. In this work, they etched back an SOI wafer to leave a 50 nm layer of silicon. Then they grew 10 nm of Si followed by 60–170 nm of $Si_{0.85}Ge_{0.15}$ by MBE with a growth temperature of 500°C. (For $Si_{0.85}Ge_{0.15}/Si$, the room temperature mismatch strain is $f = -0.0062$, corresponding to $h_c = 17$ nm.) The various thicknesses were obtained using shadow masking, to keep all other growth conditions the same. It was found that, for a 170 nm thick layer of $Si_{0.85}Ge_{0.15}$ on the ultrathin SOI substrate, lattice relaxation occurred by the introduction of dislocations in the thin Si layer rather than the $Si_{0.85}Ge_{0.15}$. After a 1 h anneal at 700°C, 800°C, or 900°C, the layer had relaxed significantly compared with the as-grown layer. Also, TEM analysis revealed that the structure annealed at 700°C contained misfit dislocations at the $Si/Si_{0.85}Ge_{0.15}$ interface, but that the associated threading segments were present only in the Si layer, not in the $Si_{0.85}Ge_{0.15}$. They interpreted these results as evidence of compliance in the Si layer, associated with slippage at the Si/SiO_2 interface. A conclusive test of compliant behavior in this material system could be made by growing various thicknesses of $Si_{0.85}Ge_{0.15}$ on an SOI layer less than $h_c = 17$ nm in thickness. Then, no misfit dislocations would be expected to form at the interface. In the work reported by Powell et al., however, the SOI layer was 60 nm thick, so the observation of interfacial dislocations does not prove a lack of compliancy in the thin Si layer.

LeGoues et al.[64] further studied the *ex situ* relaxation of SiGe on SOI compliant substrates. The SOI wafer used in this study was produced using SIMOX, and had a top Si layer 65 nm thick with a dislocation density of about $10^5\,cm^{-2}$. They grew a 180 nm thick $Si_{0.85}Ge_{0.15}$ layer

($f = -0.0062$ and $h_c = 17$ nm) by MBE at 400°C, and observed the relaxation and dislocation structure in the as-grown sample and after thermal annealing. In the as-grown sample, no dislocations were observed by cross-sectional TEM inspection, indicating a threading dislocation density of less than the resolution of the technique (10^6 cm^{-2}). Also, XRD measurements revealed that the in-plane lattice constant of the $Si_{0.85}Ge_{0.15}$ matched that of the underlying Si. The as-grown $Si_{0.85}Ge_{0.15}$ was therefore believed to be pseudomorphic. Upon annealing at 700°C or 900°C in an inert ambient, the $Si_{0.85}Ge_{0.15}$ relaxed by the formation of 60° dislocations at the interface. However, the associated threading dislocations were observed only in the thin Si layer and not in the $Si_{0.85}Ge_{0.15}$. They interpreted these results as evidence of compliancy in the thin silicon layer on the SIMOX wafer. However, they did not compare this behavior with the case of growth on standard silicon control wafers. Also, as with the previous study by Powell et al., they did not grow on a Si layer of less than 17 nm thickness to test the ability to grow a pseudomorphic layer of any thickness.

Yang et al.[62] demonstrated the growth of GaN on SiC on SOI. In their work, they produced a thin layer of SiC on a bonded and etched SOI (BESOI) wafer by exposing the top silicon layer to a flux of carbon or acetylene at 900°C. Then, they grew GaN on the SiC-on-SOI wafer using a 10 nm AlN nucleation layer, a 10 nm GaN layer, and 10 periods of AlN/GaN superlattice with 4 nm periodicity. They grew a top GaN layer with 200 nm thickness, but gave few details of its material properties.

Seaford et al.[65] compared the MBE growth of GaAs on Si(511) and SOI (511) wafers. The SOI (511) wafer was fabricated by bonding, and the top layer of silicon was thinned to 100 nm. The GaAs grown on the SOI (511) wafer had a 25% reduction in the XRD 004 full width at half maximum compared with growth on the control substrate. Also, the threading dislocation density on the SOI wafer was lower by an order of magnitude, as determined by cross-sectional TEM characterization. Here, the SOI layer was insufficiently thin to provide a conclusive test of its compliancy.

Pei et al.[66] also studied the growth of GaAs-on-SOI (511) wafers, with top silicon layers having thicknesses of 100 and 200 nm. They showed by XTEM that the GaAs on the thinner (100 nm) SOI layer had a lower threading dislocation density than the GaAs grown on the thicker (200 nm) SOI layer. Growth free from misfit dislocations could not be demonstrated on either wafer, however, because both Si layers were thicker than the critical layer thickness for the GaAs/Si heteroepitaxial system.

Despite the aforementioned published results, the question remained as to whether an SOI layer could serve as a true compliant substrate, mechanically released from its handle substrate. In an attempt to answer this question, Rehder et al.[63] made a detailed experimental and modeling study of SiGe relaxation on SOI substrates. The $Si_{0.82}Ge_{0.18}$ layers ($f = -0.0074$ and $h_c = 14$ nm) were grown by VPE to various thicknesses at temperatures of 550°C, 630°C, and 670°C. $Si_{0.82}Ge_{0.18}$ was also grown at 700°C, but only to a thickness of 6 nm because it roughened immediately. The substrates included SOI wafers with various Si thicknesses (40, 70, 330, and 10,000 nm), as well as bulk Si control wafers. The resulting samples were examined by XRD, AFM, and TEM.

Rehder et al. found that pseudomorphic $Si_{0.82}Ge_{0.18}$, 150 nm thick, could be grown on 40 or 70 nm thick SOI layers. These metastable films could be relaxed *ex situ* by annealing in the range of 875°C–1050°C. Following annealing at 950°C, the SOI developed a strain (0.047% and 0.035% for the 40 and 70 nm SOI layers, respectively). However, these values of strain were only about one-quarter of the values expected for an ideally compliant layer. In addition, the strain in the SOI layer only appeared in conjunction with the broadening of the SiGe XRD peak and the appearance of surface cross-hatch, both of which are indirect indications of misfit dislocation production at the SiGe/Si interface.

Rehder et al. also studied the *in situ* relaxation of $Si_{0.82}Ge_{0.18}$ grown at 630°C. SiGe layer thicknesses of 150, 340, 765, and 1200 nm were chosen, resulting in a wide range of *in situ* strain relaxation; the 150 nm layer is unrelaxed, whereas 80% of the mismatch strain is relaxed in the thickest layer.

In order to understand whether the SOI behaved as a compliant substrate, Rehder et al. compared their experimental results with four equilibrium models for the strain in the thin Si layer of the SOI. In the compliant substrate model of Lo,[54] if it is assumed that $\nu_{siGe}=\nu_{Si}$ and $E_{siGe}=E_{Si}$, then the in-plane strains in the SiGe and Si will be

$$\varepsilon_{SiGe} = \frac{f}{1 + h_{SiGe} / h_{Si}}, \tag{8.37}$$

and

$$\varepsilon_{Si} = \frac{-f}{1 + h_{Si} / h_{SiGe}}, \tag{8.38}$$

respectively.

Rehder et al. developed three additional models by equating the line tension on the misfit segment of a dislocation (at the SiGe/Si interface) with the strain force exerted on the threading segment of the dislocation in the thin SOI layer. In Model 1, the line tension of the misfit segment of a dislocation at the interface was assumed to be the same as in the case of growth on a thick, noncompliant substrate. Neglecting the core parameter, this is given by

$$F_L = \frac{Gb\left(1 - \nu\cos^2\alpha\right)}{4\pi(1-\nu)} \ln\left(\frac{h_{SiGe}}{b}\right), \tag{8.39}$$

where:

- G has been assumed to be equal for the epitaxial layer and the substrate
- α is the angle between the Burgers vector and the line vector for the dislocations
- h_{SiGe} is the epitaxial layer thickness

However, they recognized that the line tension of the misfit segment would be reduced by the presence of the SiO_2 layer, because of its lower shear modulus. In Model 2, the line tension was calculated using Equation 8.39, but the average shear modulus for Si and SiO_2 was used. In developing Model 3, they assumed that the oxide acts as a free surface, leading to a modified line tension given by

$$F_L = \frac{Gb\left(1 - \nu\cos^2\alpha\right)}{4\pi(1-\nu)} \ln\left(\frac{h_{Si}h_{SiGe}}{b\left(h_{Si} + h_{SiGe}\right)}\right), \tag{8.40}$$

where:

- G has been assumed to be equal for the epitaxial layer and the substrate
- α is the angle between the Burgers vector and the line vector for the dislocations

h_{SiGe} is the epitaxial layer thickness
h_{Si} is the SOI thickness

For all three models, the strain force on the threading segment of the dislocation in the SOI was calculated using

$$F_{TD} = \frac{Gb\varepsilon_{Si}h_{Si}(1+\nu)}{(1-\nu)}, \tag{8.41}$$

where ε_{Si} is the in-plane strain in the silicon layer. The equilibrium strain in the Si is predicted to be

$$\varepsilon_{Si} = \frac{b(1-\nu\cos^2\alpha)}{4\pi h_{Si}(1+\nu)}\ln\left(\frac{h_{\mathrm{SiGe}}}{b}\right) \quad \text{(Model 1)}, \tag{8.42}$$

$$\varepsilon_{Si} = \left(\frac{G_{Si}+G_{SiO_2}}{2G_{Si}}\right)\frac{b(1-\nu\cos^2\alpha)}{4\pi h_{Si}(1+\nu)}\ln\left(\frac{h_{\mathrm{SiGe}}}{b}\right) \quad \text{(Model 2)}, \tag{8.43}$$

$$\varepsilon_{Si} = \frac{b(1-\nu\cos^2\alpha)}{4\pi h_{Si}(1+\nu)}\ln\left(\frac{h_{Si}h_{\mathrm{SiGe}}}{b(h_{Si}+h_{\mathrm{SiGe}})}\right) \quad \text{(Model 3)}, \tag{8.44}$$

or

$$\varepsilon_{Si} = \frac{|f|}{1+h_{Si}/h_{\mathrm{SiGe}}} \quad \text{(Lo model)}. \tag{8.45}$$

Figure 8.37 shows the *out-of-plane strain* calculated using these four models, along with the experimental results of Rehder et al. The calculated results are shown for Models 1 (dashed curve), 2 (dotted curve), and 3 (solid curve). The solid curve at the top labeled "compliant substrate" was calculated using the Lo model. The experimental results for $Si_{0.82}Ge_{0.18}$ layers with thicknesses of 1200 and 765 nm are plotted as well, and can be fit very well using Model 3. However, the strain partitioning in the SOI layers does not follow the compliant substrate theory.

In summary, Rehder et al. found that the dependence of the $Si_{0.82}Ge_{0.18}$ relaxation on the temperature and thickness was the same on bulk Si and SOI wafers. In all cases, relaxation of the $Si_{0.82}Ge_{0.18}$ was accompanied by the introduction of misfit dislocations at the SiGe/Si interface. Tensile strain in the Si, predicted by compliant substrate theory, only occurred with the introduction of interfacial misfit dislocations. Moreover, the amount of strain in the Si was too small to be attributed to a compliant substrate mechanism. The only important effect of the SOI substrate is that the buried oxide layer reduces the line energies of misfit dislocations. Whereas a compliant substrate is supposed to increase the critical layer thickness for an epitaxial overlayer, the reduction in the misfit dislocation line energy actually

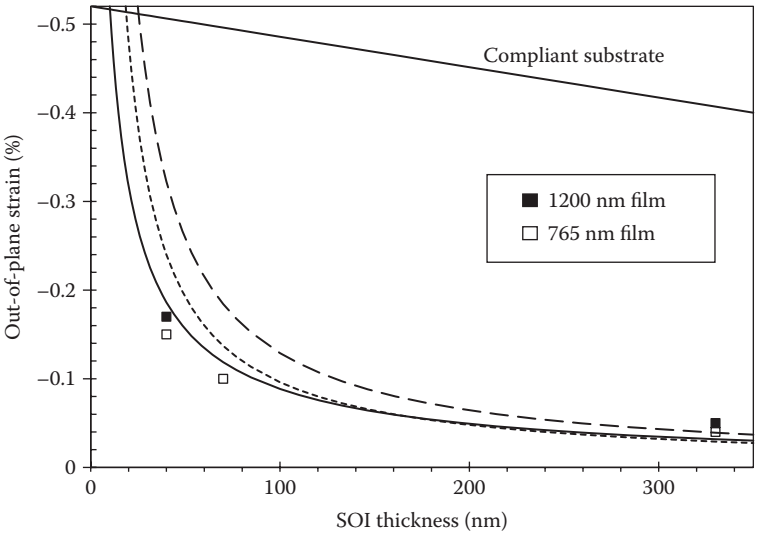


FIGURE 8.37
Out-of-plane strain in an SOI layer as a function of the Si thickness. The calculated results are shown for Models 1 (dashed curve), 2 (dotted curve), and 3 (lower solid curve), and the Lo compliant substrate model (upper solid curve). Also shown are experimental results for the growth of $Si_{0.82}Ge_{0.18}$ layers on SOI substrates, with $Si_{0.82}Ge_{0.18}$ thicknesses of 1200 and 765 nm. (Adapted from E. M. Rehder et al., *J. Appl. Phys.*, 94, 7892 [2003]. With permission. Copyright 2003, American Institute of Physics.)

decreases the critical layer thickness. These results show that, in the work of Rehder et al., the SOI did not behave as a compliant substrate for the overgrowth of SiGe.

8.9.2.3 Twist-Bonded Compliant Substrates

Ejeckam et al.[67,68] invented an approach involving the twist bonding of two wafers, followed by the thinning of the top wafer to render it compliant. They called the twist-bonded structure a compliant universal (CU) substrate.[69]

The fabrication process developed by Ejeckam et al.[68] for a twist-bonded compliant substrate is shown schematically in Figure 8.38. The process begins with two standard GaAs(001) wafers (Figure 8.38a). An AlAs etch stop layer and a 100 Å thick compliant layer of GaAs are grown epitaxially on one of the wafers. Next, the two wafers are bonded together with a twist angle. The top GaAs substrate is etched away to the etch stop layer, and then the AlAs layer itself is removed by another selective etch step. This leaves only the thin (compliant) GaAs layer twist bonded to the bottom wafer.

At the twist boundary, there is a large angular misalignment (~10°) between the ⟨110⟩ directions of the compliant layer and the substrate; however, the ⟨001⟩ directions are parallel. The result is a dense square array of screw dislocations, with spacing d given by Frank's rule:

$$d = \frac{b}{2\sin(\theta/2)}, \tag{8.46}$$

where:

- b is the length of the Burgers vector
- θ is the twist angle

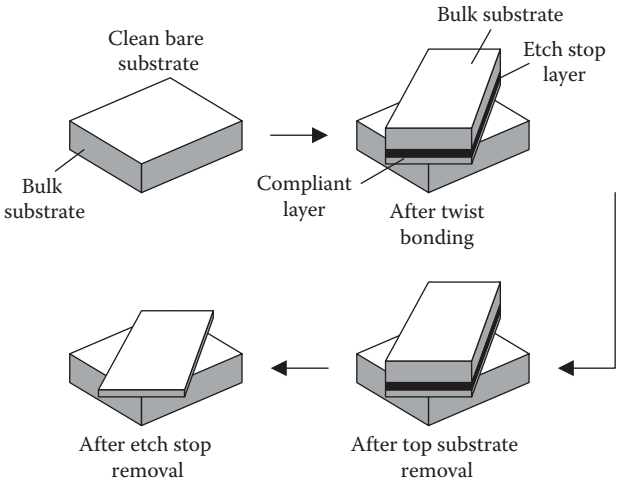


FIGURE 8.38
Fabrication process for a twist-bonded compliant substrate. (Adapted from F. E. Ejeckam et al., *Appl. Phys. Lett.*, 70, 1685 [1997]. With permission. Copyright 1997, American Institute of Physics.)

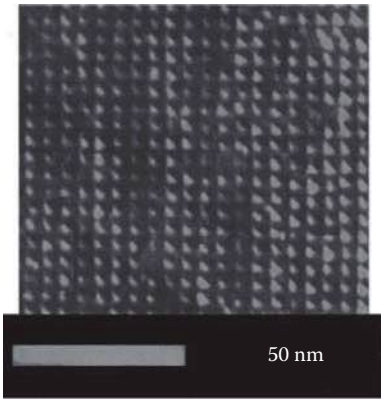


FIGURE 8.39
Plan-view dark-field weak-beam TEM micrograph of a twist boundary created by bonding a 100 Å GaAs compliant layer to a GaAs(001) substrate. The twist angle is 4.2°, and the spacing of the screw dislocations is $d = 5.3$ nm. (Reprinted from F. E. Ejeckam et al., *Appl. Phys. Lett.*, 70, 1685 [1997]. With permission. Copyright 1997, American Institute of Physics.)

Figure 8.39 shows a plan-view TEM micrograph of such a twist boundary created by bonding a 10 nm GaAs compliant layer to a GaAs(001) substrate. The twist angle is 4.2°. The spacing of the screw dislocations is $d = 5.3$ nm, which is very close to the value predicted by Frank's rule ($d = 5.5$ nm).

The atomic structure of the twist boundary is shown schematically in Figure 8.40 for the case of simple cubic crystals. The open circles represent atoms in the thin, compliant layer, while the closed circles represent atoms in the substrate wafer. Inside the square regions, the atoms in the twist-bonded layer line up with the atoms in the underlying substrate. But in the boundaries between the square regions, the atoms in the compliant layer are displaced significantly by the array of screw dislocations.

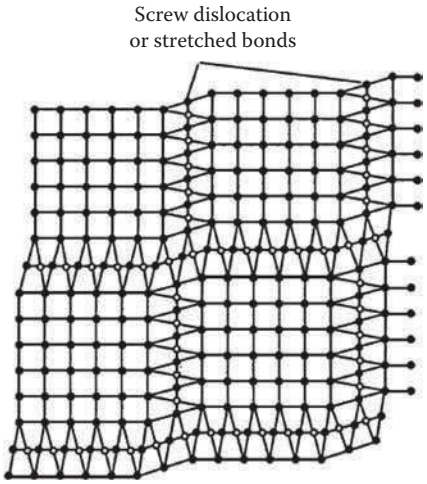


FIGURE 8.40
Schematic of a twist boundary between simple cubic crystals. The open circles represent atoms in the thin, compliant layer, while the closed circles represent atoms in the substrate wafer. (Reprinted from F. E. Ejeckam et al., *Appl. Phys. Lett.*, 70, 1685 [1997]. With permission. Copyright 1997, American Institute of Physics.)

Jesser et al.[70] made a detailed study of the implementation of twist-bonded compliant substrates. Two of their key findings were that the twist angle should be large (greater than about 8°) and that coincidence angles should be avoided. A large twist angle results in overlapping strain fields for the screw dislocations at the boundary, rendering the thin layer more ideally compliant. On the other hand, a coincidence angle (one that causes a large number of lattice sites to align on either side of the boundary) should be avoided because this locks the thin twist-bonded layer into a deep energy minimum with respect to the handle wafer, and renders it noncompliant.

The design requirements for a twist-bonded compliant substrate, as enumerated by Jesser et al.,[70] are summarized below:

1. The compliant layer should be as thin as possible, but not so thin that the screw dislocations at the twist boundary are attracted to its surface.
2. The twist angle should be greater than about 8°.
3. Coincidence angles should be avoided.
4. Ideally, the compliant layer should be selected to have a small lattice mismatch with the heteroepitaxial material that will be grown on top of it, but the compliant layer need not be lattice matched with the handle substrate.
5. The compliant layer should be chosen to induce layer-by-layer growth of the heteroepitaxial layer on top of it; island growth will lead to geometrically necessary dislocations where the islands coalesce.
6. Ideally, the compliant layer should have a smaller Young's modulus than the heteroepitaxial material that will be grown on top of it.
7. The handle wafer should have a large mismatch with respect to the compliant layer, achieved through either a large twist angle or a large lattice constant mismatch, but need not be of the same crystal structure as the compliant layer.

Practical twist-bonded compliant substrates often satisfy several, but not all, of these design criteria. Nonetheless, twist-bonded GaAs compliant substrates have been used with varying levels of success for the growth of heteroepitaxial InGaP, $In_{0.22}Ga_{0.78}As$, GaSb, and InSb.

Ejeckam et al.[68] used twist-bonded GaAs compliant substrates to grow $In_{0.35}Ga_{0.65}P$. Their compliant substrates included a 100 Å top layer bonded to a GaAs substrate with a twist angle of 9°, 17°, or 32°. They grew 300 nm of $In_{0.35}Ga_{0.65}P$ on the twist-bonded compliant substrates by MOVPE. For this composition, $f = 0.01$ and $h_c \approx 10$ nm. Though these layers are ~30 times the critical layer thickness, they were found to be free from dislocations for all three values of the twist angle investigated.

In another study, Ejeckam et al.[69] used twist-bonded GaAs compliant substrates to grow InSb ($f = -12.7\%$). The twist-bonded wafer had a 3 nm compliant layer bonded with a twist angle of $40° \pm 5°$. On these compliant substrates, they were able to grow pseudomorphic layers of InSb up to 650 nm thick, many times the critical layer thickness.

Other bonding approaches have also been tried in order to create compliant substrates. For example, Doolittle et al.[71] demonstrated a compliant substrate technology in which a GaN thin film was grown epitaxially on lithium gallate (LGO), $LiGaO_2$, removed by selective etching, and then bonded on GaAs. Glass-bonded wafers have also been investigated for use as compliant substrates.[60,61] Moran et al. studied the growth of $In_{0.44}Ga_{0.56}As$ on GaAs compliant substrates that were glass-bonded using either 10% B_2O_3 or 30% B_2O_3 borosilicate glass. The 10% and 30% borosilicate glasses have viscosities of 10^{17} P and 10^{12} P, respectively, at the growth temperature of 650°C. $In_{0.44}Ga_{0.56}As$ layers 3 μm thick were grown on conventional GaAs substrates, 12° twist-bonded compliant substrates, and both low- and high-viscosity glass-bonded compliant substrates. They found that the material grown on the glass-bonded compliant substrates (either low or high viscosity) had the best crystal quality (as judged by the 004 x-ray rocking curve width). Interestingly, the material grown on the twist-bonded wafer appeared to be *inferior* to that grown on the conventional GaAs substrate.

8.10 Free-Standing Semiconductor Films

Due to the difficulties associated with the bulk growth of GaN substrates, III-nitride devices are implemented exclusively using heteroepitaxy on mismatched substrates at the present time. This approach, while successful, gives rise to large threading dislocation densities (10^8–10^{10} cm^{-2}) and also significant thermal strains due to the mismatched thermal expansion coefficients of the III-nitrides and their sapphire or SiC substrates. An alternative approach is the realization of free-standing GaN by its heteroepitaxy and subsequent removal from the substrate.

Kelly et al.[72] first demonstrated free-standing GaN films by growth on sapphire and laser-induced liftoff. They first grew 250–300 μm thick GaN on a 2 inch *c*-face sapphire substrate by hydride vapor phase epitaxy (HVPE). Then, with the heteroepitaxial structure heated to 600°C, a pulsed laser beam was scanned over the sample surface. The resulting decomposition of a thin layer of GaN near the interface caused the liftoff of the GaN film. Nearly complete liftoff was achieved, and the resulting GaN film had an area almost equal to that of the starting sapphire substrate.

Tsai et al.[73] produced free-standing GaN films with EPDs less than 4×10^4 cm^{-2}. In this work, a 2 μm thick GaN template layer was first grown on a sapphire (0001) substrate by MOVPE. Next, a thick layer (50–200 μm) of GaN was grown on the template by HVPE.

The resulting thick film of GaN was next separated from its sapphire substrate using laser-induced liftoff. Finally, an additional thickness of GaN was grown by HVPE. For the characterization of template layers and free-standing films, H_3PO_4:H_2SO_4 was used as the crystallographic etch, and the resulting etch pits were observed by AFM. In 2 μm thick MOVPE template layers, the EPDs were found to be as high as 6×10^8 cm^{-2}. In a 500 μm thick free-standing GaN film, however, no etch pits were observed in the 50 μm × 50 μm field of view, therefore placing an upper limit of 4×10^4 cm^{-2} on the EPD.

While this approach appears to be promising, several problems need to be solved before it can be exploited commercially. Two critical problems are the incomplete separation of the heteroepitaxial layer from the substrate and the retention of curvature following separation.

8.11 Conclusion

A number of defect engineering approaches for heteroepitaxial layers have emerged. These include buffer layer approaches, dislocation compensation, patterned growth, dislocation sidewall getering, selective evaporation of defects, ELO, NHE, and compliant substrates, to name a few. All were designed to reduce the dislocation densities of heteroepitaxial layers to practical levels for device applications. Some are intended to remove existing defects from lattice-relaxed heteroepitaxial layers, such as DSG, selective evaporation of defects, ELO, or superlattice buffer layers (dislocation filters). Others were conceived in order to prevent lattice relaxation in the first place; these include patterned growth, NHE, and compliant substrates.

The proliferation of defect engineering methods could be taken as an indication that none of them is uniquely suited to the purpose, for all material systems. On the other hand, some of these approaches have been highly successful, to the point of being used in commercial devices. Graded buffer layers are the most important example of this, and have been used in commercial $GaAs_{1-x}P_x$ LEDs on GaAs substrates and $In_xGa_{1-x}As$ high-electron-mobility transistors (HEMTs) on GaAs substrates. ELO is an important method used to reduce the threading dislocation densities in the active regions of III-nitride lasers. Other defect engineering approaches, such as the use of compliant substrates, show great promise, and yet their commercial exploitation is not yet in sight.

In order to tap the great potential of heteroepitaxy, defect engineering approaches will continue to be important, not only in the applications listed above, but in new ones as well. The commercial need to integrate devices from different material systems will surely drive the development of new defect engineering approaches.

PROBLEMS

1. Suppose a uniform $GaAs_{0.5}P_{05}$ layer is to be grown on a GaAs(001) substrate without any compositional grading. Determine the approximate requirement on the thickness in order to achieve a threading dislocation density of $<10^6$ cm^{-2} at the top of the layer.
2. Suppose that a $GaAs_{0.5}P_{05}$ device layer is to be grown on a GaAs(001) substrate using a linearly graded buffer layer, with the same requirement on the threading dislocation density as in Problem 1. Choose the thickness and grading coefficient for the graded layer in order to achieve this with the minimum total thickness.

3. $In_{0.35}Ga_{0.65}As$ has been grown on InP with a thickness of 2 μm. If the material is to be patterned into square mesas and annealed, what is the maximum width of the mesas such that all of the threading dislocations may be removed by glide to the sidewalls?
4. If $In_{0.05}Ga_{0.95}As$ is grown on nanopatterned GaAs(001), what is the critical layer thickness for planar heteroepitaxy? What diameter of the seed pads will cause the critical layer thickness to double relative to planar heteroepitaxy? At what diameter for the GaAs seed pads will the critical layer thickness diverge to infinity?
5. $Si_{0.9}Ge_{0.1}$ is to be grown on a truly compliant layer of Si having a thickness of 100 cm. Determine the expected strains in the epitaxial layer and compliant substrate layer if the $Si_{0.9}Ge_{0.1}$ is grown to a thickness of 200 cm. Repeat this calculation assuming that the Si layer is not compliant.

References

1. J. F. Ocampo, E. Suarez, D. Shah, P. B. Rago, F. C. Jain, and J. E. Ayers, Overshoot graded layers for mismatched heteroepitaxial devices, *J. Electron. Mater.*, 37, 1035 (2008).
2. K. Arimoto, M. Watanabe, J. Yamanaka, K. Nakagawa, K. Sawano, Y. Shiraki, N. Usami, and K. Nakajima, Strain relaxation mechanisms in compositionally uniform and step-graded SiGe films grown on Si(110) substrates, *J. Solid State Electron.*, 53, 1135, (2009).
3. S. I. Molina, F. J. Pacheco, D. Araujo, R. Garcia, A. Sacedon, E. Calleja, Z. Yang, and P. Kidd, Strain relief in linearly graded composition buffer layers: A design scheme to grow dislocation free ($<10^5$ cm^{-2}) and unstrained layers, *Appl. Phys. Lett.*, 65, 2460 (1994).
4. A. Sacedon, F. Gonzalez-Sanz, E. Calleja, E. Munoz, S. I. Molina, F. J. Pacheco, D. Araujo, et al., Design of InGaAs linear graded buffer layers, *Appl. Phys. Lett.*, 66, 3334 (1995).
5. T. Kujofsa and J. E. Ayers, Design of dislocation-compensated ZnS_ySe_{1-y}/GaAs(001) heterostructures, *J. Electron. Mater.*, 42, 3034 (2013).
6. T. Kujofsa and J. E. Ayers, Evolution of kinetically-limited lattice relaxation and threading dislocations in temperature-graded ZnSe/GaAs(001) metamorphic heterostructures, *J. Electron. Mater.*, 44, 3030 (2015).
7. P. L. Gourley, T. J. Drummond, and B. L. Doyle, Dislocation filtering in semiconductor superlattices with lattice-matched and lattice-mismatched layer materials, *Appl. Phys. Lett.*, 49, 1101 (1986).
8. T. Soga, S. Hattori, S. Sakai, M. Takeyasu, and M. Umeno, MOCVD growth of GaAs on Si substrates with AlGaP and strained superlattice layers, *Electron. Lett.*, 20, 916 (1984).
9. T. Soga, S. Hattori, S. Sakai, and M. Umeno, Epitaxial growth and material properties of GaAs on Si grown by MOCVD, *J. Cryst. Growth*, 77, 498 (1986).
10. H. Okamoto, Y. Watanabe, Y. Kadota, and Y. Ohmachi, Dislocation reduction in GaAs on Si by thermal cycles and InGaAs/GaAs strained-layer superlattices, *Jpn. J. Appl. Phys.*, 26, L1950 (1987).
11. T. S. Rao, K. Nozawa, and Y. Horikoshi, Migration enhanced epitaxy growth of GaAs on Si with $(GaAs)_{1-x}(Si_2)_x$/GaAs strained layer superlattice buffer layers, *Appl. Phys. Lett.*, 62, 154 (1993).
12. T. S. Rao, K. Nozawa, and Y. Horikoshi, Structural properties of $(GaAs)_{1-x}(Si_2)_x$ layers on GaAs(100) substrates grown by migration enhanced epitaxy, *Jpn. J. Appl. Phys.*, 30, L547 (1991).
13. T. S. Rao and Y. Horikoshi, Growth of $(GaAs)_{1-x}(Si_2)_x$ metastable alloys using migration-enhanced epitaxy, *J. Cryst. Growth*, 115, 328 (1991).
14. W. Qian, M. Skowronski, and R. Kaspi, Dislocation density reduction in GaSb films grown on GaAs substrates by molecular beam epitaxy, *J. Electrochem. Soc.*, 144, 1430 (1997).

15. J. L. Reno, S. Chadda, and K. Malloy, Dislocation density reduction in CdZnTe(100) on GaAs using strained layer superlattices, *Appl. Phys. Lett.*, 63, 1827 (1993).
16. W. Qian, M. Skowronski, M. De Graef, K. Doverspike, L. B. Rowland, and D. K. Gaskill, Microstructural characterization of α-GaN films grown on sapphire by organometallic vapor phase epitaxy, *Appl. Phys. Lett.*, 66, 1252 (1995).
17. E. Feltin, B. Beaumont, M. Laügt, P. de Mierry, P. Vennéguès, H. Lahrèche, M. Leroux, and P. Gibart, Stress control in GaN grown on silicon (111) by metalorganic vapor phase epitaxy, *Appl. Phys. Lett.*, 79, 3230 (2001).
18. W. H. Sun, J. P. Zhang, J. W. Yang, H. P. Maruska, M. A. Khan, R. Liu, and F. A. Ponce, Fine structure of AlN/AlGaN superlattice grown by pulsed atomic-layer epitaxy for dislocation filtering, *Appl. Phys. Lett.*, 87, 211915 (2005).
19. N. A. El-Masry, J. C. Tarn, and N. H. Karam, Interactions of dislocations in GaAs grown on Si substrates with InGaAs-GaAsP strained layer superlattices, *J. Appl. Phys.*, 64, 3672 (1988).
20. J. W. Matthews, S. Mader, and T. B. Light, Accommodation of misfit across the interface between crystals of semiconducting elements or compounds, *J. Appl. Phys.*, 41, 3800 (1970).
21. E. A. Fitzgerald, P. D. Kirchner, R. Proano, G. D. Pettit, J. M. Woodall, and D. G. Ast, Elimination of interface defects in mismatched epilayers by a reduction in growth area, *Appl. Phys. Lett.*, 52, 1496 (1988).
22. E. A. Fitzgerald, G. P. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall, Nucleation mechanisms and the elimination of misfit dislocations at mismatched interfaces by reduction in growth area, *J. Appl. Phys.*, 65, 2220 (1989).
23. M. Kim, P. Hashemi, and J. L. Hoyt, Increased critical thickness for high Ge-content strained SieGe-on-Si using selective epitaxial growth, *Appl. Phys. Lett.*, 97, 262106 (2010).
24. X. G. Zhang, A. Rodriguez, X. Wang, P. Li, F. C. Jain, and J. E. Ayers, Complete removal of threading dislocations from mismatched layers by patterned heteroepitaxial processing, *Appl. Phys. Lett.*, 77, 2524 (2000).
25. X. G. Zhang, P. Li, G. Zhao, D. W. Parent, F. C. Jain, and J. E. Ayers, Removal of threading dislocations from patterned heteroepitaxial semiconductors by glide to sidewalls, *J. Electron. Mater.*, 27, 1248 (1998).
26. D. Hull and D. J. Bacon, *Introduction to Dislocations*, 3rd ed. (New York: Pergamon, 1984).
27. N. Chand and S. N. G. Chu, Elimination of dark line defects in GaAs-on-Si by post-growth patterning and thermal annealing, *Appl. Phys. Lett.*, 58, 74 (1991).
28. I. Yonenaga, Dynamic behavior of dislocations in InAs:In comparison with III-V compounds and other semiconductors, *J. Appl. Phys.*, 84, 4209 (1998).
29. N. N. Ledentsov, M. V. Maximov, D. Bimberg, T. Maka, C. M. Sotomayor Torres, I. V. Kochnev, I. L. Krestnikov, et al., 1.3 μ*m* luminescence and gain from defect-free InGaAs-GaAs quantum dots grown by metal-organic chemical vapor deposition, *Semicond. Sci. Technol.*, 15, 604 (2000).
30. D. S. Sizov, M. V. Maksimov, A. F. Tsatsul'nikov, N. A. Cherkashin, N. V. Kryzhanovskaya, A. B. Zhukov, N. A. Maleev, et al., The influence of heat treatment conditions on the evaporation of defect regions in structures with InGaAs quantum dots in the GaAs matrix, *Semiconductors*, 36, 1020 (2002).
31. N. Ledentsov, Semiconductor device and method of making same, U.S. Patent 6653166 (2003).
32. V. Shchukin and N. Ledentsov, Defect-free semiconductor templates for epitaxial growth and method of making same, U.S. Patent 6784074 (2004).
33. N. N. Ledentsov, V. A. Shchukin, T. Kettler, K. Posilovic, D. Bimberg, L. Ya. Karachinsky, A. Yu. Gladyshev, et al., MBE-grown metamorphic lasers for applications at telecom wavelengths, *J. Cryst. Growth*, 301–302, 914 (2007).
34. L. Jastrzebski, SOI by CVD: Epitaxial lateral overgrowth (ELO) process-review, *J. Cryst. Growth*, 63, 493 (1983).
35. R. W. McClelland, C. O. Bozler, and J. C. C. Fan, A technique for producing epitaxial films on reusable substrates, *Appl. Phys. Lett.*, 37, 560 (1980).
36. R. P. Gale, R. W. McClelland, J. C. C. Fan, and C. O. Bozler, Lateral epitaxial overgrowth of GaAs by organometallic chemical vapor deposition, *Appl. Phys. Lett.*, 41, 545 (1982).

37. P. Vohl, C. O. Bozler, R. W. McClelland, A. Chu, and A. J. Strauss, Lateral growth of single-crystal InP over dielectric films by orientation-dependent VPE, *J. Cryst. Growth*, 56, 410 (1982).
38. M. Kastelic, I. Oh, C. G. Takoudis, J. A. Friedrich, and G. W. Neudeck, Selective epitaxial growth of silicon in pancake reactors, *Chem. Eng. Sci.*, 43, 2031 (1988).
39. J. L. Glenn Jr., G. W. Neudeck, C. K. Subramanian, and J. P. Denton, A fully planar method for creating adjacent 'self-isolating' silicon-on-insulator and epitaxial layers by epitaxial lateral overgrowth, *Appl. Phys. Lett.*, 60, 483 (1992).
40. G. Shahidi, B. Davari, Y. Taur, J. Warnock, M. R. Wordeman, P. McFarland, S. Madder, et al., Fabrication of cmOS on ultrathin SOI obtained by epitaxial lateral overgrowth and chemical-mechanical polishing, *IEDM Tech. Dig.*, 587 (1990).
41. O.-H. Nam, M. D. Bremser, T. S. Zheleva, and R. F. Davis, Lateral epitaxy of low defect density GaN layers via organometallic vapor phase epitaxy, *Appl. Phys. Lett.*, 71, 2638 (1997).
42. O.-H. Nam, M. D. Bremser, B. L. Ward, R. J. Memanich, and R. F. Davis, Growth of GaN and $Al_{0.2}Ga_{0.8}N$ on patterned substrates via organometallic vapor phase epitaxy, *Jpn. J. Appl. Phys.*, 36, L532 (1997).
43. T. S. Zheleva, O.-H. Nam, M. D. Bremser, and R. F. Davis, Dislocation density reduction via lateral epitaxy in selectively grown GaN structures, *Appl. Phys. Lett.*, 71, 2472 (1997).
44. J. Park, P. A. Grudowski, C. J. Eiting, and R. D. Dupuis, Selective-area and lateral epitaxial overgrowth of III-N materials by metal organic chemical vapor deposition, *Appl. Phys. Lett.*, 73, 333 (1998).
45. Y. S. Chang, S. Naritsuka, and T. Nishinaga, Optimization of growth condition for wide dislocation-free GaAs on Si substrate by microchannel epitaxy, *J. Cryst. Growth*, 192, 18 (1998).
46. T. Zheleva, S. A. Smith, D. B. Thomson, T. Gehrke, K. J. Linthicum, P. Rajagopal, E. Carlson, W. M. Ashmawi, and R. F. Davis, Pendeo-epitaxy—A new approach for lateral growth of gallium nitride structures, *MRS Internet J. Nitride Semicond.*, 4S1, G3.38 (1999).
47. S. Einfeldt, A. M. Roskowski, E. A. Preble, and R. F. Davis, Strain and crystallographic tilt in uncoalesced GaN layers grown by maskless pendeoepitaxy, *Appl. Phys. Lett.*, 80, 953 (2002).
48. S. Luryi and E. Suhir, New approach to the high quality epitaxial growth of lattice-mismatched materials, *Appl. Phys. Lett.*, 49, 140 (1986).
49. D. Zubia and S. D. Hersee, Nanoheteroepitaxy: The application of nanostructuring and substrate compliance to the heteroepitaxy of mismatched semiconductor materials, *J. Appl. Lett.*, 85, 6492 (1999).
50. J. P. Hirth and A. G. Evans, Damage of coherent multilayer structures by injection of dislocations or cracks, *J. Appl. Phys.*, 60, 2372 (1986).
51. D. Zubia, S. H. Zaidi, S. R. J. Brueck, and S. D. Hersee, Nanoheteroepitaxial growth of GaN on Si by organometallic vapor phase epitaxy, *Appl. Phys. Lett.*, 76, 858 (2000).
52. S. H. Zaidi, A. S. Chu, and S. R. J. Brueck, Scalable fabrication and optical characterization of nm Si structures, *Mater. Res. Soc. Symp. Proc.*, 358, 957 (1995).
53. S. H. Zaidi, A. S. Chu, and S. R. J. Brueck, Optical properties of nanoscale, one-dimensional silicon grating structures, *J. Appl. Phys.*, 80, 6997 (1996).
54. Y. H. Lo, New approach to grow pseudomorphic structures over the critical layer thickness, *Appl. Phys. Lett.*, 59, 2311 (2005).
55. D. Teng and Y. H. Lo, Dynamic model for pseudomorphic structures grown on compliant substrates: An approach to extend the critical thickness, *Appl. Phys. Lett.*, 62, 43 (1993).
56. C. L. Chua, W. Y. Hsu, C. H. Lin, G. Christensen, and Y. H. Lo, Overcoming the pseudomorphic critical layer thickness limit using compliant substrates, *Appl. Phys. Lett.*, 64, 3640 (1994).
57. C. Carter-Coman, A. S. Brown, N. M. Jokerst, D. E. Dawson, R. Bicknell-Tassius, Z. C. Feng, K. C. Rajikumar, and G. Dagnall, Strain accommodation in mismatched layers by molecular beam epitaxy: Introduction of a new compliant substrate technology, *J. Electron. Mater.*, 25, 1044 (1996).
58. C. Carter-Coman, R. Bicknell-Tassius, R. G. Benz, A. S. Brown, and N. M. Jokerst, Analysis of GaAs substrate removal etching with citric acid:H_2O_2 and NH_4OH:H_2O_2 for application to compliant substrates, *J. Electrochem. Soc.*, 144, L29 (1997).

59. P. D. Moran, D. M. Hansen, R. J. Matyi, J. M. Redwing, and T. F. Kuech, Realization and characterization of ultrathin GaAs-on-insulator structures, *J. Electrochem. Soc.*, 146, 3506 (1999).
60. P. D. Moran, D. M. Hansen, R. J. Matyi, J. G. Cederberg, L. J. Mawst, and T. F. Kuech, InGaAs heteroepitaxy on GaAs compliant substrates: X-ray diffraction evidence of enhanced relaxation and improved structural quality, *Appl. Phys. Lett.*, 75, 1559 (1999).
61. A. R. Powell, S. S. Iyer, and F. K. LeGoues, New approach to the growth of low dislocation relaxed SiGe material, *Appl. Phys. Lett.*, 64, 1856 (1994).
62. Z. Yang, F. Guarin, I. W. Tao, W. I. Wang, and S. S. Iyer, Approach to obtain high quality GaN on Si and SiC-on-silicon-on-insulator compliant substrate by molecular-beam epitaxy, *J. Vac. Sci. Technol. B*, 13, 789 (1995).
63. E. M. Rehder, C. K. Inoki, T. S. Kuan, and T. F. Kuech, SiGe relaxation on silicon-on-insulator substrates: An experimental and modeling study, *J. Appl. Phys.*, 94, 7892 (2003).
64. F. K. LeGoues, A. Powell, and S. S. Iyer, Relaxation of SiGe thin films grown on Si/SiO_2 substrates, *J. Appl. Phys.*, 75, 7240 (1994).
65. M. L. Seaford, D. H. Tomich, K. G. Eyink, L. Grazulis, K. Mahalingham, Z. Yang, and W. I. Wang, Comparison of GaAs grown on standard Si(511) and compliant SOI (511), *J. Electron. Mater.*, 29, 906 (2000).
66. C. W. Pei, J. B. Héroux, J. Sweet, W. I. Wang, J. Chen, and M. F. Chang, High quality GaAs grown on Si-on-insulator compliant substrates, *J. Vac. Sci. Technol. B*, 20, 1196 (2002).
67. F. E. Ejeckam, Y. H. Lo, M. Seaford, H. Q. Hou, and B. E. Hammons, presented at Proceedings of the 9th Annual Meeting of the IEEE Lasers and Electro-Optics Society (IEEE/LEOS), Boston, (November 18–21, 1996).
68. F. E. Ejeckam, Y. H. Lo, S. Subramanian, H. Q. Hou, and B. E. Hammons, Lattice engineered compliant substrate for defect-free heteroepitaxial growth, *Appl. Phys. Lett.*, 70, 1685 (1997).
69. F. E. Ejeckam, M. L. Seaford, Y.-H. Lo, H. Q. Hou, and B. E. Hammons, Dislocation-free InSb grown on GaAs compliant universal substrates, *Appl. Phys. Lett.*, 71, 776 (1997).
70. W. A. Jesser, J. H. van der Merwe, and P. M. Stoop, Misfit accommodation by compliant substrates, *J. Appl. Phys.*, 85, 2129 (1999).
71. W. A. Doolittle, T. Kropewnicki, C. Carter-Coman, S. Stock, P. Kohl, N. M. Jokerst, R. A. Metzger, et al., Growth of GaN on lithium gallate substrates for development of a GaN thin compliant substrate, *J. Vac. Sci. Technol. B*, 16, 1300 (1998).
72. M. K. Kelly, R. P. Vaudo, V. M. Phanse, L. Gorgens, O. Ambacher, and M. Stutzmann, Large free-standing GaN substrates by hydride vapor phase epitaxy and laser-induced liftoff, *Jpn. J. Appl. Phys.*, Part 2, 38, L217 (1999).
73. C.-C. Tsai, C.-S. Chang, and T.-Y. Chen, Low-etch-pit-density GaN substrates by regrowth on free-standing GaN films, *Appl. Phys. Lett.*, 80, 3718 (2002).

9

Metamorphic Devices

9.1 Introduction

The metamorphic approach is often preferred for device fabrication because it allows considerable freedom in choosing layer compositions and thicknesses when compared with pseudomorphic growth. An important example of this is the $In_xGa_{1-x}As/In_yAl_{1-y}As$-based high-electron-mobility transistor (HEMT) grown on a GaAs substrate. The indium content in the quantum well must be limited to less than 20% if pseudomorphic growth is to be achieved on GaAs substrates, while use of InP substrates allows lattice matching with ~50% indium content. On the other hand, $In_xGa_{1-x}As/In_yAl_{1-y}As$ HEMTs with virtually any indium composition may be fabricated on GaAs substrates using metamorphic buffer layers, avoiding the need for InP substrates and extending the performance of the resulting transistors. The same is true for HEMTs based on other material systems grown metamorphically on GaAs substrates. It is still desirable to grow the quantum wells coherently strained to the confining layer material, but this is far less restrictive when using lattice-relaxed metamorphic buffer layers that can adopt any lattice constant within a wide range spanning from GaP(0.5451 nm) to GaSb(0.6096 nm).

For the development of metamorphic devices, the two most important issues are the threading dislocation densities and the roughening of the surfaces and interfaces. Much of the development of metamorphic layers has focused on HEMTs because majority carrier devices are less sensitive to dislocations in their active regions. However, the roughening of the quantum well interfaces in metamorphic HEMTs can negatively impact carrier mobility and device performance. Minority carrier devices, including heterojunction bipolar transistors (HBTs), light-emitting diodes (LEDs), laser diodes (LDs), and photovoltaics have also been developed using metamorphic structures, but their performance attributes are often limited by threading dislocations. For example, quantum well lasers can exhibit premature failure as a consequence of the movement and multiplication of dark-line defects associated with threading dislocations. On the other hand, metamorphic quantum dot lasers are less susceptible to these effects and have been fabricated with good performance and reliability.

The following sections describe some of the important applications of metamorphic materials to devices, including strained-channel metal oxide semiconductor field effect transistors (MOSFETs) on strain-relaxed buffers (SRBs), HEMTs, HBTs, LEDs, and photovoltaic cells. This treatment is not meant to be exhaustive, but illustrates many of the considerations involved in metamorphic device design.

9.2 Strain-Relaxed Buffer MOSFETs

Theoretical and experimental work has shown that SRB MOSFETs should offer significant performance improvements over standard Si devices in deep submicron complementary metal oxide semiconductor circuitry.[1] For n-channel devices, it is advantageous to employ tensile-strained silicon, while for p-channel devices, improved transport can be achieved in compressively strained SiGe channels.[2] This not only improves the carrier transport (low-field mobility and overshoot velocity) in both types of devices, but also provides a better balance between the n- and p-channel transistors. Both types of devices can be realized on an SRB of appropriate composition. The n-channel device will take on a tensile strain as long as it has lower germanium content than the buffer, while the p-channel device will be compressively strained if it has higher germanium content than the buffer. A typical application of this approach involves a buffer composition of $Si_{0.7}Ge_{0.3}$ with a silicon n-channel transistor and a $Si_{0.3}Ge_{0.7}$ p-channel transistor, as shown in Figure 9.1.

O'Neill and Antonaidis[1] have shown by a modeling study that SRB MOSFETs of the types shown in Figure 9.1 can exhibit significantly improved performance over conventional silicon n-channel and p-channel MOSFETs. Figures 9.2 and 9.3 show the cutoff frequency f_T versus the gate voltage for n-channel and p-channel SRB MOSFETs, respectively, with 100 nm gate lengths in both cases. Also shown are the cutoff frequencies for conventional Si MOSFETs. The SRB MOSFETs exhibit f_T improvements of 50% (n-channel) and 100% (p-channel) compared with conventional silicon devices.

The successful implementation of SRB MOSFETs requires that the buffer be strain relaxed while the channel layer is pseudomorphically strained. A key challenge in the

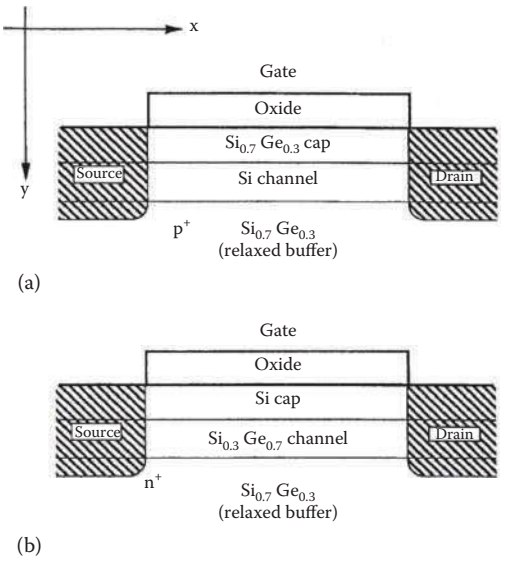


FIGURE 9.1
Strained-channel MOSFETs on SRB layers. (a) n-Channel SRB MOSFET with a tensile-strained Si channel on a relaxed $Si_{0.7}Ge_{0.3}$ buffer. (b) p-Channel SRB MOSFET with a compressively strained $Si_{0.3}Ge_{0.7}$ channel on a relaxed $Si_{0.7}Ge_{0.3}$ buffer. Both devices utilize a capping layer to reduce the effect of carrier scattering associated with slow charge states at the oxide interface. (Reprinted from A. G. O'Neill and D. A. Antoniadis, *IEEE Trans. Electron. Dev.*, 43, 911 [1996]. With permission. Copyright 1996, Institute of Electrical and Electronics Engineers.)

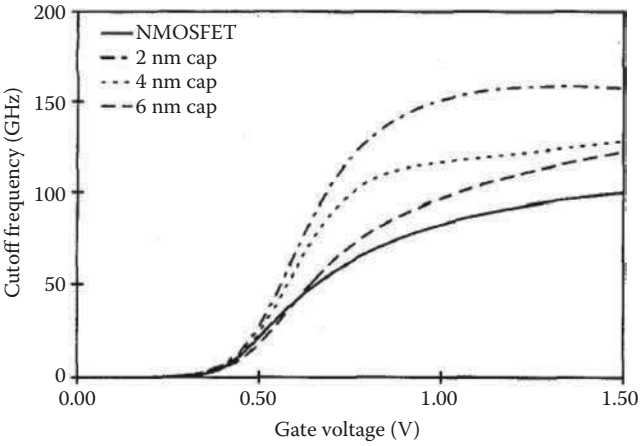


FIGURE 9.2
Cutoff frequency f_T as a function of gate voltage for n-channel SRB MOSFETs with a Si channel and a $Si_{0.7}Ge_{0.3}$ cap, grown on an SRB of $Si_{0.3}Ge_{0.7}$. Also shown is the cutoff frequency for a comparable Si n-channel MOSFET (solid curve). (Reprinted from A. G. O'Neill and D. A. Antoniadis, *IEEE Trans. Electron. Dev.*, 43, 911 [1996]. With permission. Copyright 1996, Institute of Electrical and Electronics Engineers.)

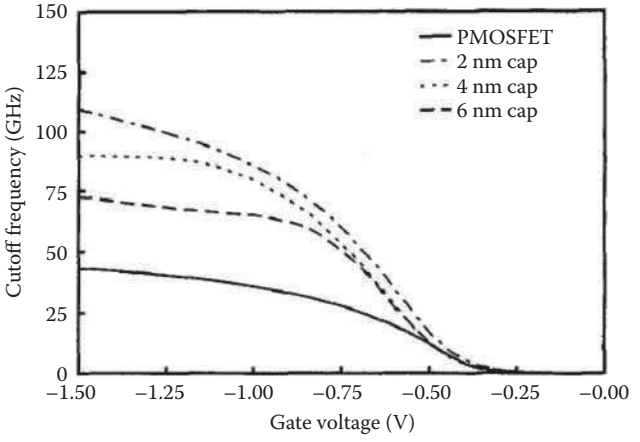


FIGURE 9.3
Cutoff frequency f_T as a function of gate voltage for p-channel SRB MOSFETs with a $Si_{0.7}Ge_{0.3}$ channel and a Si cap, grown on an SRB of $Si_{0.3}Ge_{0.7}$. Also shown is the cutoff frequency for a comparable Si p-channel MOSFET (solid curve). (Reprinted from A. G. O'Neill and D. A. Antoniadis, *IEEE Trans. Electron. Dev.*, 43, 911 [1996]. With permission. Copyright 1996, Institute of Electrical and Electronics Engineers.)

practical implementation of these devices has been the high threading dislocation density that accompanies strain relaxation in the buffer layer. LeGoues et al.[3] have shown that the use of a thick buffer layer, with a grading rate of 10% Ge/μm, enables the achievement of low threading dislocation densities of $\sim 10^5$ cm^{-2} with a top germanium content of 30%. However, the successful implementation of SRB MOSFETs in deeply scaled, high-density circuits may require the use of much thinner buffer layers, thereby necessitating the use of novel compositional grading schemes or defect engineering approaches.

9.3 High-Electron-Mobility Transistors

A number of metamorphic HEMT designs have been demonstrated using various material systems and buffer layer approaches. Devices with arsenide-based channels are the most studied, and have been demonstrated using a diversity of buffer layer approaches, including uniform composition buffers,[4] step-graded buffers,[5–10] and linearly graded buffers.[11–23] Nonlinearly graded buffers[21] and hybrid approaches involving superlattices along with graded buffers[9,23,24] have also been applied to metamorphic HEMTs with arsenide-based channels. HEMTs with nitride-based channels have also been studied extensively for high-power and high-frequency applications, because of their combination of high breakdown field and high carrier velocity. These devices have typically been realized on 6H-SiC(0001) or sapphire (0001) substrates using GaN buffer layers.

9.3.1 HEMTs with Arsenide Channel Layers

The first metamorphic $In_xGa_{1-x}As/In_xAl_{1-x}As$ HEMT to be realized on a GaAs substrate in the range $0.25 \leq x \leq 0.50$ was demonstrated by Grider et al.[5] using a step-graded $In_xGa_{1-x}As$ buffer grown at low temperature (300°C) by molecular beam epitaxy (MBE). They observed that dislocations were confined to the low-temperature, metamorphic buffer using cross-sectional transmission electron microscopy (XTEM). Metamorphic HEMTs fabricated using the low-temperature step-graded buffer exhibited 77K Hall mobilities in the range 22,000–27,450 cm^2/Vs (with a maximum corresponding to 40% indium) and 300K Hall mobilities in the range of 5120–6880 cm^2/Vs. However, use of a similar step-graded buffer deposited at higher temperature (500°C) gave rise to a greater density of threading dislocations observable by XTEM, as well as lower Hall mobilities (7200 and 4240 cm^2/Vs at 77K and 300K, respectively). The superior Hall mobilities achieved with low-temperature buffers were attributed to the reduced threading dislocation densities, although the role of interfacial roughness was not studied.

Some studies of metamorphic HEMTs suggest that the buffer layer surface smoothness exerts more influence than the threading dislocation density in determining the two-dimensional electron gas (2DEG) mobility. The cross-hatch and corrugations associated with metamorphic growth can give rise to significant electron scattering and lower the 2DEG mobility if the interface roughness has a lateral scale similar to the Fermi wavelength, given by $\lambda_F = \sqrt{2\pi / n_s}$, where n_s is the sheet electron concentration.[25] Thomas et al.[26] grew $InAs/Al_xGa_{1-x}Sb$ quantum wells on GaAs substrates using uniform 2 μm thick uniform buffers of either GaSb or AlSb by MBE. Both types of structures contained ~10^8 cm^{-2} threading dislocations, but the one with the GaSb buffer exhibited much higher 2DEG mobility at 12K: 944,000 cm^2/Vs with $n_s = 1.3 \times 10^{12}$ cm^{-2} compared with 244,000 cm^2/Vs with the AlSb buffer, both for a 2DEG concentration of $n_s = 1.3 \times 10^{12}$ cm^{-2}. This difference was attributed to the smoother morphology of the structure with the GaSb buffer due to the higher surface mobility of Ga adatoms. However, it is possible that the relative roles played by threading dislocations and surface roughness depend on the material system.

Recent work with metamorphic HEMTs has emphasized $In_xGa_{1-x}As/In_yAl_{1-y}As$ quantum wells grown on GaAs(001) substrates by MBE.[10,11,13–16,18,19,22,27] Hybrid buffer approaches have also been used, such as a linearly graded buffer, along with a superlattice.[12,17,23] Several comparative studies have been made in an attempt to evaluate different buffer layer approaches. Song et al.[28] compared HEMT devices on linearly graded layers with different grading coefficients and doping. They found that moderate grading gave a lower

threading dislocation density than either lower or higher grading coefficients, and that a Be-doped buffer was superior to a Si-doped buffer in terms of defect density. Cordier et al.[16] also studied the effects of grading coefficient and buffer layer thickness, for $In_{0.33}Ga_{0.67}As$/$In_{0.34}Al_{0.66}As$ HEMTs on linearly graded $In_xAl_{1-x}As$ with and without overshoot. The inclusion of overshoot resulted in lower residual strain in the device layers, but the strain relaxation, as well as the surface roughness, 77K electron mobility, and crystallographic tilt, changed considerably as the buffer thickness, and therefore grading coefficient, was varied.

Some studies have compared the use of $In_xGa_{1-x}As$ and $In_xAl_yGa_{1-y}As$ linearly graded buffers in metamorphic HEMTs. Inoue et al.[11] found that use of the wider bandgap buffer containing aluminum resulted in a lower residual carrier concentration, and higher resistivity, by an order of magnitude. Chyi et al.[29] when comparing linearly graded $In_xGa_{1-x}As$ and $In_xAl_{1-x}As$ buffers, found that the aluminum-containing buffer relaxed less completely, and exhibited less crystallographic tilt, presumably because of differences in dislocation mobility.

Choi et al.[21] compared HEMTs on sublinear, linear, and superlinear $In_xAl_{1-x}As$ buffer layers. They found that the sublinear profile gave the lowest threading dislocation density (1.3×10^8 cm^{-2}), the superlinear profile gave the highest threading dislocation density (2.6×10^8 cm^{-2}), and the linearly graded device was intermediate. This may indicate the importance of the misfit dislocation free zone (MDFZ) adjacent to the device in reducing the defect density.

Some work has considered the use of a superlattice along with a linearly graded buffer to improve the quality of the device layers. Galiev et al. compared a device utilizing a linearly graded layer with overshoot with a similar transistor having a linearly graded buffer plus a superlattice.[23] In their work, inclusion of the superlattice resulted in a lower dislocation density and smoother surface. Mendach et al.[24] and Heyn et al.[9] showed that insertion of a superlattice buffer lowered the threading dislocation density in metamorphic HEMTs with linearly graded buffers and InAs channels.

Modern metamorphic HEMTs on GaAs substrates exhibit performance and reliability comparable with pseudomorphic HEMTs fabricated on InP wafers, but with a wider available range of indium mole fraction in the channel layers. Wakita et al.[14] showed that HEMTs, with a channel indium mole fraction of 0.53, grown on GaAs using a low-temperature linearly graded buffer exhibit low-bias reliability similar to p-HEMTs on InP with an extrapolated mean time to failure (MTTF) of 7.5×10^6 h at 125°C and an activation energy of 1.7 eV.

9.3.2 HEMTs with Nitride Channel Layers

HEMTs utilizing nitride-based channels are of great interest for radio frequency and power switching applications because of several material advantages compared with GaAs- or InP-based transistors. First, the high breakdown strength (~3×10^6 V/cm) allows the use of higher drain voltages than in AlGaAs/GaAs HEMTs. Second, the large conduction band offset at the AlGaN/GaN interface and strain-induced piezoelectric charge gives rise to high sheet carrier densities in the 2DEG of up to 10^{13} cm^{-2}. Third, the saturated electron drift velocity is relatively high, on the order of 10^7 cm/s. These properties are expected to allow high output power densities in the X- and Ka-bands.

Most nitride-based HEMT devices have been fabricated using metamorphic structures on lattice-mismatched substrates, including 4H-SiC(0001),[30–42] sapphire (0001),[31,43–49] and Si(111).[50–54] Silicon carbide substrates have been preferred by many researchers because of their high thermal conductivity (~3.3 W/cmK), which is about seven times that of sapphire,

enabling high power density without the need for flip-chip mounting schemes.[30] However, SiC wafer technology is immature compared with that of silicon, so that the starting wafers are of smaller diameter, higher cost, and lower quality.[50] This has led to the use of Si(111) substrates, which offer a reasonable thermal conductivity (~1.3 W/cmK) between that of sapphire and 4H-SiC, but at lower cost than either. Significant challenges in the growth of GaN on Si are the large lattice mismatch and the considerable difference in thermal expansion coefficients. The latter problem gives rise to significant thermal strain with the possibility of cracking in the heteroepitaxial layers. The use of mismatched substrates (Si, SiC, or sapphire) has generally led to high threading dislocation densities of the order of 10^8 to 10^9 cm^{-2}, serving as the motivation for the fabrication of nitride-based HEMTs on free-standing GaN substrates.[55]

Most nitride-based HEMTs have utilized an AlGaN/GaN heterostructure, with a typical device design shown in Figure 9.4. Although this figure shows a device on a sapphire substrate, realization on 4H-SiC(0001) or Si(111) is similar. A thin (~20 nm), low-temperature AlN nucleation layer is used, followed by the growth of a thick (~3 μm) GaN buffer layer to reduce the threading dislocation density. An $Al_xGa_{1-x}N$ supply layer is grown on top to establish a 2DEG in the adjacent GaN, with typical aluminum mole fractions varying from 0.14 to 0.40. The $Al_xGa_{1-x}N$ may be donor doped with silicon, but it has been established that most of the sheet carrier density in the 2DEG results from the strong piezoelectric effect in the nitrides.[56] In doped devices, an undoped $Al_xGa_{1-x}N$ spacer (~3 nm) is used to reduce carrier scattering in the channel by ionized impurities in the supply layer[46] and an undoped cap layer is used to form a Schottky gate. Most reported devices were grown by metalorganic vapor phase epitaxy (MOVPE), but some researchers have also applied MBE.[34,36,37,52] Multilayer ohmic contacts to the source and drain are formed by the deposition of Ti/Al/Ni/Au[44] or Ti/Al/Ti/Au,[31,43,50] followed by thermal annealing. The gate is formed by Ni/Au[32,43,55] or Pt/Au,[44] and typically multilayer resist schemes may be used to create a mushroom cross section.[43] He or Ne ion implantation can be used to isolate devices, as shown in Figure 9.4. Although the HEMT is inherently a depletion-type device, enhancement mode operation has been demonstrated by using a fluoride-based plasma treatment prior to the gate formation.[46] Normally-off operation has also been achieved using an InGaN cap layer.[48] Short-channel effects may be suppressed by using a double heterostructure design utilizing a back barrier of the high-bandgap material.[41]

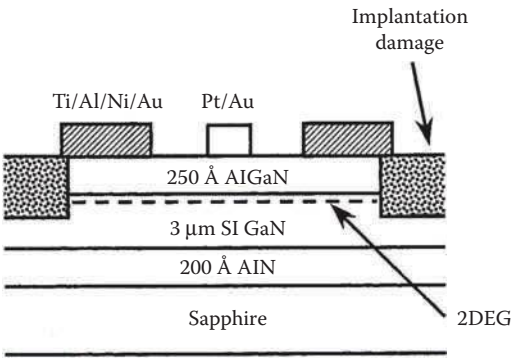


FIGURE 9.4
Typical structure for AlGaN/GaN HEMT on a sapphire substrate. (Reprinted from S. Binari et al., *IEEE Trans. Electron Dev.*, 48, 465 [2001]. With permission. Copyright 2001, Institute of Electrical and Electronics Engineers.)

One important area of research for nitride-based HEMTs involves the surface passivation of slow-acting trapping states between the drain and the gate, which are believed to be responsible for the frequency-dependent current degradation of the devices.[43] These trapping states may be associated with dangling bonds, the high density of threading dislocations, and possibly surface contamination. Surface passivation has been studied using several approaches, including the plasma-enhanced chemical vapor deposition of Si_3N_4,[43] the thermal or plasma-enhanced atomic layer deposition of Al_2O_3,[42] or the atomic layer deposition of AlN.[54] Huang et al.[54] showed that surface passivation by AlN was effective in suppressing current collapse and reducing the dynamic on resistance of AlGaN/GaN HEMTs. However, it has also been shown that the effects of the surface states may be diminished by using a thick GaN cap layer, in the absence of passivation.[33]

Although most nitride-based HEMTs have utilized AlGaN/GaN structures, these suffer from buffer leakage and short-channel effects. In order to minimize these effects, several research groups have implemented double heterostructures or InGaN channel layers to achieve better carrier confinement. Two potential drawbacks of using an InGaN channel are indium clustering and alloy scattering, which tend to reduce the 2DEG mobility.[47] However, Okamoto et al.[45] demonstrated an AlGaN/InGaN/GaN double heterostructure HEMT with a channel mobility of >1000 cm^2/Vs with a sheet carrier density of 1.3×10^{13} cm^{-2}. Liu et al.[47] showed that, by using an InGaN "notch" layer between the GaN channel layer and the GaN buffer, they could achieve better confinement and high 2DEG mobility (~1300 cm^2/Vs) because of the relative lack of alloy scattering compared with the case of using an InGaN channel.

Excellent performance has been demonstrated in nitride-based HEMTs despite the high threading dislocation densities associated with metamorphic growth on mismatched substrates. Lee et al.[41] demonstrated an InAlN/GaN device on a SiC substrate with a current gain cutoff frequency of $f_T = 300$ GHz using a channel length of 30 nm and an InGaN back barrier to suppress short-channel effects. Shinohara et al.[39] reported a double-heterostructure AlGaN/GaN device on a SiC substrate with $f_T = 220$ GHz and $f_{max} = 400$ GHz using a 40 nm channel length. Crespo et al.[38] demonstrated high-power Ka-band performance of an InAlN/GaN HEMT on a SiC substrate, with 5.8 W/mm output power at 35 GHz with a power added efficiency of 43.6%.

9.4 Heterojunction Bipolar Transistors

Metamorphic HBTs[22,57,58] and double-heterojunction bipolar transistors (DHBTs)[59] have been fabricated on GaAs and Si substrates using linearly graded buffers. Yang et al.[57] compared the thermal resistances of InP-based HBTs on GaAs substrates using $In_xAl_{1-x}As$ and $In_xGa_{1-x}P$ graded buffers by an experimental and modeling study. They found that the commonly used linearly graded $In_xAl_{1-x}As$ buffer contributed significantly to the device thermal resistance, and accounted for 40% of the total thermal resistance in the case of a 250 nm thick buffer. An $In_xGa_{1-x}P$ buffer of similar thickness contributed much less to the thermal resistance, rendering it about 20% more than for a pseudomorphic device on an InP substrate. Lew et al.[58] demonstrated an InGaP/GaAs HBT on a Si substrate using a linearly graded $Si_{1-x}Ge_x$ buffer, with a maximum current gain of ~25 at a collector current of 100 mA and an offset voltage of 150 mV. Tsai[22] reported the integration of InP/InGaAs HBTs and InGaAs/InP HEMTs on a GaAs substrate using a linearly graded $In_xGa_{1-x}P$

buffer layer. The HBT exhibited a maximum current gain of 255 with a low offset voltage of 105 mV.

9.5 Light-Emitting Diodes

LEDs have been fabricated using a wide range of epitaxial materials and substrates, but recent work has emphasized III-phosphide and III-nitride devices, and state-of-the-art high-power LEDs are made exclusively from these two material systems. With few exceptions, LEDs are all metamorphic so that issues of strain, strain-induced roughening, stacking faults, misfit dislocations, and threading dislocations have profoundly affected the course of their development.

The visible spectrum extends from approximately 400 nm (3.1 eV) to 700 nm (1.8 eV). Figure 9.5 shows the luminous efficiency of the human eye, with the approximate color regions labeled. The relationship between the wavelength and photon energy is given by[60]

$$\lambda = \frac{hc}{E_g} = \frac{1240\ \text{nm/eV}}{E_g}, \tag{9.1}$$

where:

- h is the Planck constant
- c is the speed of light

For a particular LED, an important figure of merit is the external quantum efficiency, which gives the ratio of the number of generated photons to the number of electrons

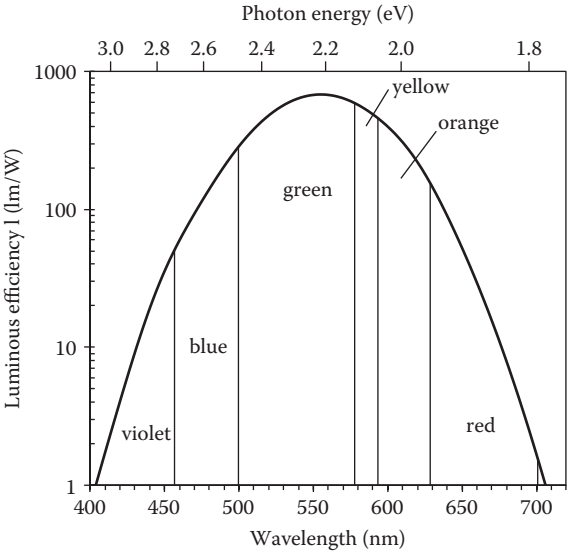


FIGURE 9.5
Approximate luminous efficiency of the human eye as a function of wavelength.

entering the LED contact. The external quantum efficiency is given by the product of the internal quantum efficiency η_{int} and the light extraction efficiency C_{ext},[61]

$$\eta_{ext} = \eta_{int} C_{ext}. \tag{9.2}$$

In practical devices, the injection efficiency can be close to unity so that the internal quantum efficiency is determined by the rates of radiative and nonradiative recombination:[62]

$$\eta_{int} = \frac{R_r}{R_r + R_{nr}}, \tag{9.3}$$

where:

R_r is the rate of radiative recombination
R_{nr} is the rate of nonradiative recombination

The radiative recombination rate is proportional to the product of the carrier concentrations,

$$R_r = Bnp, \tag{9.4}$$

where B is a constant and n and p are the electron and hole concentrations, respectively, and the nonradiative recombination rate is given by the Schockley–Read–Hall equation,

$$R_{nr} = \frac{np - n_i^2}{n\tau_h + p\tau_e}, \tag{9.5}$$

where:

n_i is the intrinsic carrier concentration
τ_e is the electron capture time
τ_h is the hole capture time

The presence of defects such as threading dislocations in the active region of the LED greatly increases the nonradiative recombination rate to the detriment of the internal quantum efficiency. Therefore, there has been much effort to fabricate LEDs on lattice-matched substrates whenever possible, or to control the threading dislocation densities through compositional grading or other defect engineering approaches when lattice matching is not possible. Even when high internal quantum efficiency is achieved, efficient extraction of light from the device is a challenge because the LED materials have high indices of refraction ($n \sim 3.5$ for III-phosphides and $n \sim 2.4$ for III-nitrides) that correspond to relatively narrow escape cones. Light that is totally internally reflected tends to be absorbed in the device structure, and this is a particularly severe problem with III-phosphides that are grown on an absorbing GaAs substrate. Therefore, much effort has gone into special physical designs to allow efficient light extraction, as well as removal of the absorbing GaAs substrate from III-phosphide devices.

Figure 9.6 shows the external quantum efficiencies for high-power III-nitride and III-phosphide LEDs as functions of the wavelength. $(Al_xGa_{1-x})_{0.52}In_{0.48}P$ LEDs allow direct bandgap emission from $x = 0$ ($E_g \sim 1.9$ eV and $\lambda \sim 650$ nm, deep red) to $x = 0.58$ ($E_g \sim 2.25$ eV and $\lambda \sim 550$ nm, yellow green). As the direct–indirect transition is approached, the internal quantum efficiency falls off dramatically, as shown in Figure 9.6, because of the

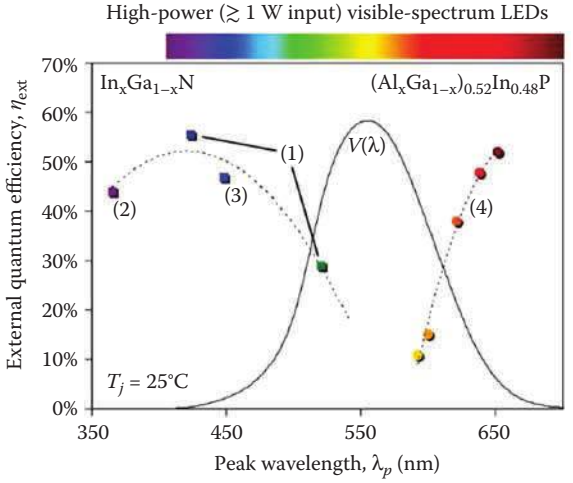


FIGURE 9.6
External quantum efficiencies of high-power LEDs fabricated in the III-nitride and III-phosphide material systems as functions of the wavelength. (Reprinted from M. R. Krames et al., *J. Display Technol.*, 3, 160 [2007]. With permission. Copyright 2007, Institute of Electrical and Electronics Engineers.)

transfer of carriers to the satellite valley in the conduction band and also due to poor carrier confinement. $In_xGa_{1-x}N$ is direct gap over the entire compositional range from $x = 0$ ($E_g \sim 3.4$ eV and $\lambda \sim 365$ nm, ultraviolet A [UV-A]) to $x = 1.0$ ($E_g \sim 0.7$ eV and $\lambda \sim 1800$ nm, infrared). However, practical considerations associated with spontaneous polarization and miscibility issues have limited the useful wavelength range of $In_xGa_{1-x}N$ LEDs from $\lambda \sim 365$ nm (UV-A) to $\lambda \sim 550$ nm (deep green).

The following sections are organized by the color of emission, with sections on red, amber, green, blue, ultraviolet, and white LEDs. Some material systems therefore appear in more than one section, but the compositions of the materials, and the important practical issues associated with them, vary along with the emission wavelength.

9.5.1 Red Light-Emitting Diodes

Red LEDs have been fabricated using a number of material systems, including GaAsP, AlGaAs, AlInGaP, InGaN, and GaNP. The red portion of the spectrum extends approximately from 620 to 700 nm, corresponding to bandgaps in the range 1.8–2.0 eV.

Early red LEDs utilized $GaAs_{1-x}P_x$ grown heteroepitaxially on GaAs(001) or GaP(001) substrates by vapor phase epitaxy in the Ga-HCl-AsH_3-PH_3 system. GaP substrates, though of lower quality, are transparent in the wavelength range of interest and allow for efficient light extraction from an inverted device structure.[63] $GaAs_{1-x}P_x$ is a direct gap semiconductor for phosphorus compositions up to ~0.40, and this alloy composition has been used in commercial red LEDs with an emission wavelength of 650 nm.[63] For higher compositions, the indirect bandgap requires doping by nitrogen, an isoelectronic trap that introduces an electronic level close to the conduction band and greatly enhances the luminescence.[64] However, even in red LEDS nitrogen doping has been used with a resulting shift of the emission to longer wavelengths and a broadening of the emission spectrum. For red LEDs utilizing an active layer alloy composition of 0.40, the lattice mismatch is 1.4% with GaAs and −2.1% with GaP. Thick, linearly graded GaAsP buffer layers have been used to accommodate this mismatch,

with typical grading rates of ~1% phosphorus/μm and total thicknesses of ~40 μm. (The associated grading constant C_β, in lattice mismatch per cm, is 3.5 cm^{-1}.) Despite the slow rates of grading utilized in these devices, (001) wafers exhibit visible cross-hatch on the surface associated with the orthogonal arrays of misfit dislocations in the graded region.[65] It is well established that the threading dislocation density in the device layer is directly proportional to the rate of grading in the buffer layer,[66] and that higher threading dislocation densities reduce the luminescence efficiency.[65,67,68] Indeed, early studies of linearly graded layers such as those by Abrahams et al.[66] were motivated by the need to control the strain and threading dislocations in these structures for the development of efficient and reliable red LEDs.

$In_xGa_{1-x}P$ LEDS were developed as an alternative to GaAsP devices due to the possibility of high brightness, as well as the ability to extend the wavelengths from 650 nm (red) to 565 nm (green). This material exhibits a direct bandgap up to an indium mole fraction of 0.37, at which the energy bandgap is $E_g = 2.17$ eV.[69] Alloys in this system have been grown by a variety of epitaxial methods, including hydride epitaxy, chloride epitaxy, liquid phase epitaxy, MOVPE, and MBE, and LEDs have been demonstrated on Si,[70] GaAs,[71] and GaP[72] substrates. LEDs on Si and GaP substrates exhibit a high degree of lattice mismatch, requiring the incorporation of thick graded buffer layers of $In_xGa_{1-x}P$. In fact, even though InGaP could be lattice matched to GaAs, the matched composition (~0.50) is not a direct gap material. Therefore, InGaP LEDs on GaAs must utilize mismatched compositions so that mismatch strain and the associated defects have important effects on device performance. For diodes on GaP substrates, the relative radiometric intensity falls off rapidly for wavelengths longer than 610 nm (orange) because of the increased mismatch and higher threading dislocation density in the active region of the device.[72] This is consistent with the observation that, on GaP substrates, the 1/*e* fall time for electroluminescence from InGaP LEDs decreases from 12 ns for green emitters to 2 ns for red emitters,[72] and this could indicate the shorter nonradiative lifetime in red LEDs with higher dislocation densities.

Red LEDs in the material system $Al_xGa_{1-x}As$ were made feasible by the development of appropriate epitaxial growth technologies, including liquid phase epitaxy (temperature difference method[73]) and MOVPE. These devices are grown on GaAs(001) substrates, and exhibit very little lattice mismatch over the entire compositional range (maximum mismatch 0.1%). Therefore, very low threading dislocation densities are obtained without the need for compositionally graded buffers, resulting in higher brightness compared with the earlier GaAsP diodes. $Al_xGa_{1-x}As$ exhibits a room temperature direct bandgap of

$$E_g(x) = \left(1.424 + x1.427 + x^2 0.041\right) \text{eV} \tag{9.6}$$

up to compositions of 0.40, beyond which the material exhibits an indirect bandgap.[74] Beyond this composition, the emission efficiency drops off dramatically, as shown in Figure 9.7. A typical single heterojunction red LED is fabricated on a p-GaAs:Zn substrate with a p-$Al_{0.35}Ga_{0.65}As$ active layer and an n-$Al_{0.65}Ga_{0.35}As$ emitter layer.[73] The active layer emits at ~650 nm, and the top layer is reasonably transparent to this wavelength because of its larger, indirect bandgap. Even though these red LEDs are largely absent the issues associated with lattice-mismatched heteroepitaxy, they are quite limited in their applicability due to the inability to tune the wavelength shorter than about 650 nm.

The capability to grow semiconductors containing both aluminum and phosphorus, brought about by the development of MOVPE, spurned great interest in the quaternary AlInGaP for high-brightness LEDS, such as the ones used in automotive applications. In particular, the quaternary alloy $(Al_xGa_{1-x})_{0.52}In_{0.48}P$ allows tuning of the energy gap from green to red by the adjustment of x, while the lattice constant remains closely matched to

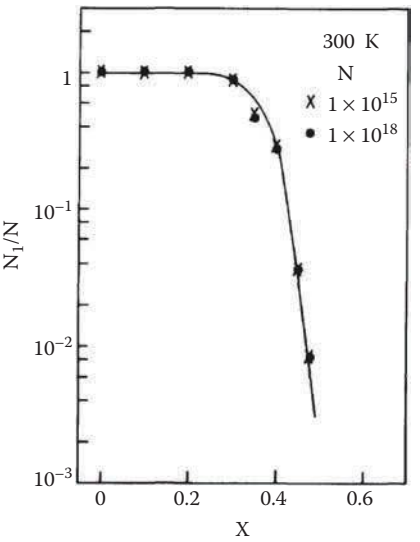


FIGURE 9.7
Emission efficiency as a function of composition x for $Al_xGa_{1-x}As$. (Reprinted from J. Nishizawa et al., *J. Appl. Phys.*, 54, 2807 [1983]. With permission. Copyright 1983, American Institute of Physics.)

GaAs substrates over the entire range. (Approximate lattice matching is obtained with 50% indium, and often this material is referred to as $(Al_xGa_{1-x})_{0.5}In_{0.5}P$.) Devices are commonly grown on misoriented GaAs substrates in order to avoid ordering of the alloy.

A typical high-brightness red $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ LED structure is shown in Figure 9.8. GaAs(001) is used almost exclusively for the substrate due to the need for lattice matching, even though it is strongly absorbing in the wavelength region of interest; that is, the need for a low threading dislocation density in the active regions outweighs the problem of an absorbing substrate and the countermeasures needed to minimize its effect. One such countermeasure is a distributed Bragg reflector (DBR), which can recover a portion of the light scattered toward the substrate. These DBRs can be fabricated from $Al_xGa_{1-x}As$-$Al_yGa_{1-y}As$ or $(Al_xGa_{1-x})_{0.5}In_{0.5}P$-$Al_{0.5}In_{0.5}P$ multilayers, and in either case approximate lattice matching to GaAs is preserved. In a double-heterostructure design such as that shown in Figure 9.8, the n-type and p-type confinement layers are typically $Al_{0.5}In_{0.5}P$:Te and $Al_{0.5}In_{0.5}P$:Mg or $Al_{0.5}In_{0.5}P$:Zn, respectively. The undoped setback layers are included to minimize unintended diffusion of the n- or p-type impurities into the active layer of $(Al_xGa_{1-x})_{0.52}In_{0.48}P$, whose composition is chosen near x = 0 for red output. The p window layer is made of transparent $Al_{0.5}In_{0.5}P$:Zn and facilitates current spreading from the top electrical contact.

Although the use of lattice-matched GaAs has been emphasized in the development of AlInGaP LEDs, Kwon et al. demonstrated red AlInGaP resonant cavity LEDs on Si substrates using a SiGe graded buffer layer.[75] In this demonstration, devices on the Ge/SiGe/Si substrates achieved higher output power than comparable LEDs on GaAs or Ge substrates due to enhanced lateral current spreading. The dislocation density in the virtual Ge/SiGe/Si substrate was reported to be 10^6 cm^{-2} due to the slow rate of grading (10% germanium/μm) and was low enough so that it did not limit the device performance.

Chip design of red LEDs is of great importance if high extraction efficiency is to be obtained. The index of refraction for $(Al_xGa_{1-x})_{0.52}In_{0.48}P$ is $n \sim 3.5$, and the critical angle for total internal reflection is given by

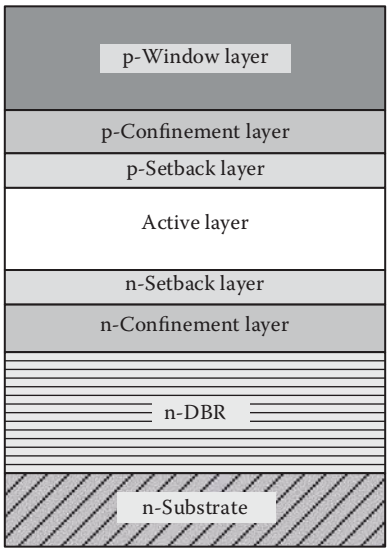


FIGURE 9.8
Typical structure of a high-brightness red $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ LED. (Reprinted from K. Streubel et al., *IEEE J. Sel. Top. Quant. Electron.*, 8, 321 [2002]. With permission. Copyright 2002, Institute of Electrical and Electronics Engineers.)

$$\theta_c = \sin^{-1}\left(\frac{n_a}{n_s}\right), \tag{9.7}$$

where

n_s is index of refraction for the semiconductor
n_a is the index of refraction for the surrounding ambient (air or epoxy)

For emission from a flat surface into air, the light extraction efficiency is only about 2%, corresponding to an escape cone with $\theta_c = 18°$. Using epoxy resin, this number can be doubled to 4% with a larger critical angle of $\theta_c = 27°$. Some of this emission is blocked by the top electrical contact so the actual extraction efficiencies will be less, but this effect can be reduced by using a current spreading layer under the top contact, as shown in Figure 9.8. Sometimes an n-type blocking layer is used under the contact,[76] to prevent carrier injection in this region and thereby avoid wasted light emission.

Several basic chip designs have emerged, as shown in Figure 9.9. The thin epitaxial structure on an absorbing substrate (thin AS, Figure 9.9a) allows a maximum light extraction efficiency of 4%, as described above. Here, the absorbing substrate is the GaAs(001) wafer used for the growth of the epitaxial structure. Use of a thick window layer on an absorbing substrate (thick AS, Figure 9.9b) allows access to one-half of each of the side-facing escape cones, giving a theoretical light extraction efficiency up to 12%. Here, the window layer must be electrically conducting but transparent to the emitted light. It does not participate in the light emission process, and so in general it can be a lattice-mismatched layer with relatively high dislocation density. Both lattice-matched (AlGaAs) and -mismatched (GaP) window layers have been used. Inclusion of a DBR at the bottom of the structure (thick As + DBR, Figure 9.9c) further increases the limiting light extraction efficiency to 16% by recovering the downward-facing escape cone, which would otherwise be absorbed by

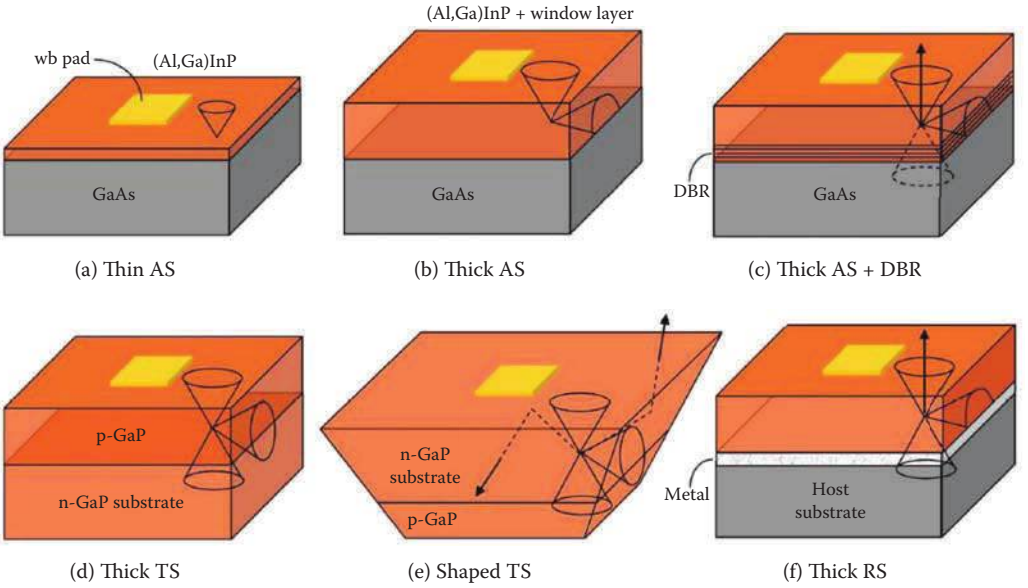


FIGURE 9.9
Chip designs for high-brightness $(Al_xGa_{1-x})_{0.52}In_{0.48}P$ LEDs. (a) Thin epitaxial structure on an absorbing substrate (AS) (b) Thick window layer on an absorbing substrate. (c) Thick window on an absorbing substrate with a DBR. (d) Thick window layer on a transparent substrate (TS). (e) Shaped window layer on a TS. (f) Thick window layer on a reflecting substrate. (Reprinted from M. R. Krames et al., *J. Display Technol.*, 3, 160 [2007]. With permission. Copyright 2007, Institute of Electrical and Electronics Engineers.)

the GaAs substrate. Further improvements in C_{ext} require the elimination of the absorbing GaAs substrate. For example, the LED may be removed from the GaAs substrate and wafer bonded to a transparent GaP substrate (thick TS, Figure 9.9d). If a reflective bottom contact is used, C_{ext} up to 24% is possible. Shaping of the transparent substrate (shaped TS, Figure 9.9e) can help harness light from outside of the normal escape cone for a rectangular parallelepiped structure. Theoretical structures that are of interest include a hemispherical dome and the compound parabolic concentrator, but these present considerable practical problems in their fabrication using semiconductor crystals, which tend to facet. A practical chip design that draws on this concept is the truncated inverted pyramid (TIP) LED design shown in Figure 9.9e, and external quantum efficiencies up to ~60% have been demonstrated. Still another approach involves using a thick window layer with a reflecting substrate (thick RS, Figure 9.9f). Here, the LED structure is bonded to a metallized wafer that acts as a reflecting substrate. Shaping and texturing of the LED surface allow improved light extraction efficiency, but this design still does not match the performance of the TIP LED.

9.5.2 Amber Light-Emitting Diodes

Early yellow LEDs were realized with GaAsP:N grown on GaAs using thick, linearly graded buffer layers. However, these devices were limited to a few percent in external quantum efficiency due to the high threading dislocation densities and poor carrier confinement. Modern high-brightness amber LEDs are realized in $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ using double heterostructures or multiple quantum well (MQW) designs of the type employed

in red LEDs. The direct (Γ) and indirect (X) bandgaps of $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ at room temperature are given approximately by

$$E_{g\Gamma}(x) = (1.900 + 0.61x)\ \text{eV} \tag{9.8}$$

and

$$E_{gX}(x) = (2.204 + 0.085x)\ \text{eV} \tag{9.9}$$

so that the crossover to an indirect gap semiconductor occurs at $x = 0.58$, where $E_g = 2.25$ eV and $\lambda = 550$ nm (green). As the photon energy is increased toward the crossover point, the internal quantum efficiency falls off rapidly due to the transfer of electrons from the Γ valley to the X valley in the conduction band. Nonetheless, high brightness may be obtained down to a wavelength of ~580 nm, which is in the amber portion of the spectrum.

In contrast to red $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ LEDs, carrier leakage is the primary loss mechanism involved in determining the internal quantum efficiency. Typically, $Al_{0.5}In_{0.5}P$ confinement layers are used, either in a double heterostructure or an MQW design, and both band offsets decrease with increasing x (increasing photon energy).

As with red LEDs, it is necessary to utilize GaAs substrates for lattice-matched growth in order to avoid high threading dislocation densities. Therefore, similar methods are used to mitigate the effect of the strongly absorbing substrate, including DBRs, substrate removal and bonding to a TS (GaP), or thin-film LED processing. All of the chip designs illustrated in Figure 9.9 are therefore applicable here as well.

9.5.3 Green Light-Emitting Diodes

Green LEDs present a significant challenge in the attainment of high external quantum efficiencies, because the green portion of the spectrum lies between the optimum wavelength ranges for InGaN- and AlInGaP-based devices. Nonetheless, there is great impetus for developing efficient, high-brightness green LEDs for red-green-blue (RGB) displays, projectors, and white solid-state lighting, and a number of materials have been used for green LEDs, including GaP, CdZnSe, AlInGaP, and InGaN.

Early green LEDs were based on GaP,[77] which is an indirect gap semiconductor. These green LEDs were grown by liquid phase epitaxy on a substrate of opposite conductivity type to create a p-n junction and emit at ~555 nm. The homoepitaxial growth avoids problems of dislocations and stacking faults, but the indirect bandgap and lack of a quantum well to confine the injected carriers leads to poor brightness and quantum efficiency. Doping by the isoelectronic impurity nitrogen introduces an electronic level close to the conduction band and greatly enhances the luminescence,[64] but the resulting performance (0.1 cd luminous intensity, 0.1% external quantum efficiency) is still below that of modern III-nitride devices.

Green LEDs were subsequently developed using II-VI heterostructures on ZnSe(001) or GaAs(001) substrates. Eason et al.[78] reported the MBE homoepitaxial growth of a green LED utilizing a single 100 nm $ZnTe_{0.1}Se_{0.9}$ quantum well between n-type and p-type ZnSe confining layers. The emission wavelength was 512 nm, and the device produced an output of 1.3 mW with an external quantum efficiency of 5.3%. Although Eason et al. utilized a lattice-matched ZnSe substrate, most work involving II-VI heterostructure green LEDs involved the use of GaAs(001) substrates because of their large area, lower cost, and high quality. However, heteroepitaxial growth of II-VI semiconductors on GaAs generally

results in high densities of stacking faults and threading dislocations, both of which have deleterious effects on device reliability and lifetime. Hovinen et al.[79] reported that the stacking faults were the original source of dark-line defects in blue-green emitters based on ZnCdSe/ZnSSe quantum wells. Moreover, it is difficult to dope these materials p-type. One approach to circumvent the doping problem is to use a blue InGaN LED to pump an undoped CdMgZnSe color-converting heterostructure. In a study of these devices, Zhu et al.[80] found that even small amounts of residual strain contributed to the formation of stacking faults and dark-line defects.

Both common LED material systems, AlInGaP and InGaN, have been used for the fabrication of single quantum well (SQW) or MQW green LEDs. However, AlInGaP-based LEDs are limited to emission wavelengths longer than about 550 nm and exhibit poor external quantum efficiencies in the green portion of the spectrum. Nakamura et al.[81] demonstrated MOVPE growth of SQW InGaN green LEDs emitting at 520 nm with high brightness (12 cd) and external quantum efficiency (6.3%). There are two difficulties in the fabrication of green InGaN LEDs associated with the use of (0001) sapphire or SiC substrates. The first involves the difficulty of incorporating indium, and the second is that the polarization-induced field gives rise to a blue shift of the emission with increasing injection current. For these reasons, some research has been conducted toward the achievement of InGaN LEDs on semipolar or nonpolar crystal orientations.[82]

When combining red, green, and blue LEDs in multicolor displays, the range of possible colors is dictated by the triangle enclosed by the three LED points plotted on a chromaticity diagram of the type shown in Figure 9.10. This diagram shows the range of achievable colors using a blue InGaN LED, a red AlGaAs LED, and several possible green LEDs, which have been demonstrated in the material systems GaP, AlInGaP, and InGaN.

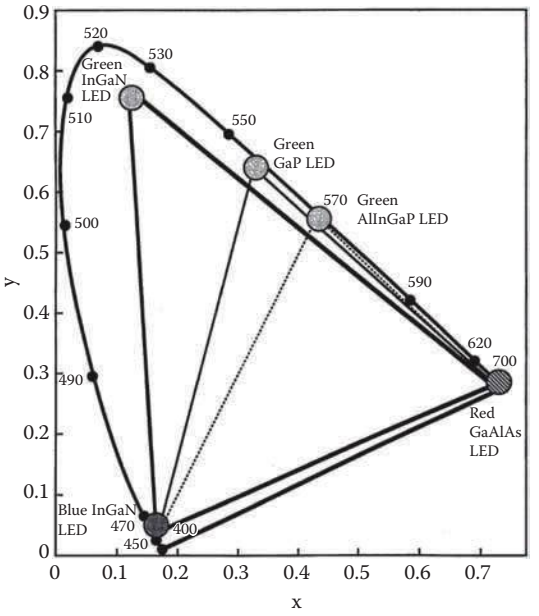


FIGURE 9.10
Chromaticity diagram showing a blue SQW InGaN LED, a red AlGaAs LED, and green LEDs fabricated using GaP, AlInGaP, and InGaN. (Reprinted from S. Nakamura et al., *Jpn. J. Appl. Phys.*, 34, L1332 [1995]. With permission. Copyright 1995, Japan Society of Applied Physics.)

Although II-VI-based emitters allow similar chromaticity performance, they suffer from issues of reliability due to the high threading dislocation densities and stacking fault densities coupled with the sensitivity of these materials to defects.

One of the challenges of InGaN-based green LEDs is that of "efficiency droop."[83] Efficiency droop is a decrease in internal quantum efficiency, occurring as current density increases.[84] Although efficiency droop is not completely understood, it has been associated with causes such as Auger recombination, poor carrier confinement, and poor efficiency of hole injection into quantum wells.[83]

Structural methods such as electron blocking layers (EBLs), superlattices, and chirp barriers have all been studied as ways to reduce the causes of efficiency droop.[83] In 2013, Kang et al. developed an AlGaN graded superlattice (GSL) for green LED performance enhancement.[83] The mechanism by which the GSL improves efficiency droop is through a reduction in hole barrier height, allowing for better efficiency of hole injection.[83] The implementation of the GSL yielded an efficiency droop of 19.4% at a current density of 100 A/cm^2.[83]

Another issue facing this technology is that InGaN-based LEDs are more susceptible to threading dislocation defects as the indium concentration is increased. This is of importance in green InGaN LEDs, which require the highest indium mole fractions in the active regions. Improved performance is made possible by defect engineering approaches such as epitaxial lateral overgrowth (ELO) or epitaxial growth on lattice-matched GaN substrates. Considerable progress has been made in the development of GaN substrates with low dislocation densities,[85] and this should benefit InGaN-based LEDs.

9.5.4 Blue Light-Emitting Diodes

The blue part of the spectrum spans from ~400 nm (~3.1 eV) to ~450 nm (~2.7 eV). The first commercial blue LEDs were based on 6H-SiC(0001) because this was the only wide-bandgap semiconductor that could be readily doped n-type and p-type. SiC is an indirect gap semiconductor, however, so the resulting devices exhibited very low efficiencies (0.03%). Typically, 6H-SiC blue LEDS utilized donor–acceptor pair transitions from an active layer doped with both Al and N, with emission at a wavelength of 470 nm.

Several breakthroughs in GaN technology, including the epitaxial growth on Al_2O_3 substrates[86] and postgrowth annealing to activate Mg acceptors,[87] allowed the realization of efficient blue LEDs in 1992. However, large, high-quality GaN wafers are not available, so III-nitride devices are nearly always grown on lattice-mismatched substrates, either SiC or Al_2O_3 (sapphire). The resulting threading dislocation densities are extremely high, typically 10^8 to 10^9 cm^{-2}.

Modern blue LEDs utilize multiple InGaN quantum wells sandwiched between AlGaN barrier layers as the active region. n-GaN and p-GaN layers inject electrons and holes into the active region, respectively. A typical structure using a sapphire substrate is shown in Figure 9.11. Because sapphire is insulating, both contacts need to be made to the top of the device when this substrate is used.

Chip design is very important if high extraction efficiencies are to be achieved, but some new issues come into play with nitride-based devices. In the case of 6H-SiC substrates, the substrate is both transparent and highly conductive, so that the thick TS and the shaped TS approaches shown in Figure 9.9 are applicable. Extraction efficiencies up to 47% have been obtained with the shaped TS design. Al_2O_3 (sapphire) is transparent but electrically

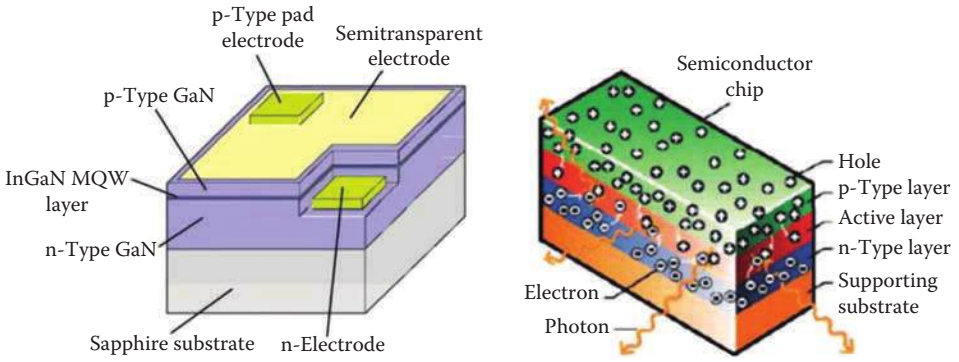


FIGURE 9.11
Typical structure for a blue III-nitride LED on a sapphire substrate. (Reprinted from S. Nakamura, *MRS Bull.*, 34, 101 [2009]. With permission. Copyright 2009, Materials Research Society.)

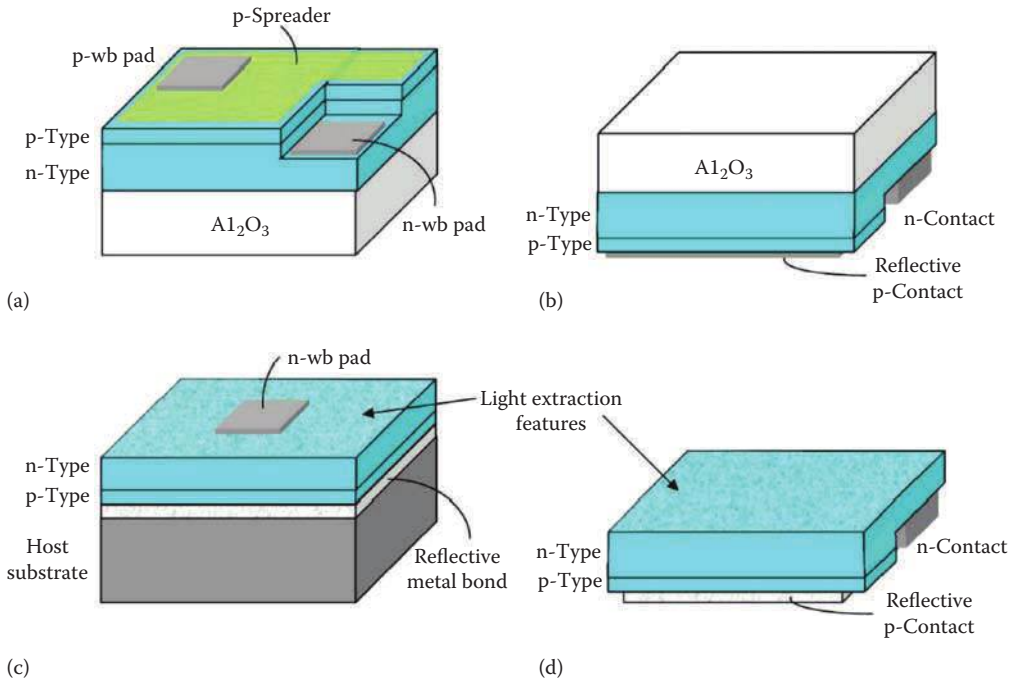


FIGURE 9.12
Chip designs for III-nitride LEDs grown epitaxially on sapphire substrates. (a) "Epi-up" conventional chip (CC) with semitransparent p-contact on the top. (b) Flip-chip (FC) with reflective p-contact on the bottom. (c) VTF bonded to reflective metal on a host substrate. (d) TFFC. (Reprinted from M. R. Krames et al., *J. Display Technol.*, 3, 160 [2007]. With permission. Copyright 2007, Institute of Electrical and Electronics Engineers.)

insulating, so that new approaches to chip design are needed in order to facilitate both electrical contacts to the diode. Some of these are illustrated in Figure 9.12.

In the conventional chip (CC) design shown in Figure 9.12a, both contacts must be made from the top of the wafer. n-Type GaN can be doped to much lower resistivity than p-type GaN, so the only viable approach is to etch down to the n-type material for its ohmic contact while making a large-area contact to the p-type material with a semitransparent

p-spreader layer typically made of NiAu or indium tin oxide (ITO). Design of the p-spreader layer involves a trade-off between current spreading and transparency. If the device is to be operated at high current density, the p-spreader layer must be made thicker, but this blocks more of the emitted light. In the flip-chip design of Figure 9.12b, this trade-off is avoided because both contacts are on the bottom. The p-contact is made directly under the active region, and so it is made from thick, highly reflective metallization such as Ag. This results in optical cavity effects that have a strong influence on the radiation pattern and the extraction efficiency. The vertical-injection thin-film (VTF) design of Figure 9.12c requires removal of the substrate, followed by bonding to a metallized host substrate wafer. The n-type GaN:Si can have a resistivity as low as 0.005 Ω cm, allowing a very thin n-type current spreading layer. Texturing of the n-type surface allows extraction efficiencies up to 75%. The thin-film flip-chip (TFFC) of Figure 9.12 also allows texturing of the n-type GaN after removal of the substrate, but avoids top-side wire bonding and has excellent bottom-side heat removal. Extraction efficiencies up to 80% have been achieved with the TFFC design, and high-power operation is possible due to the good thermal characteristics.

As with green InGaN LEDs, efficiency droop is an important problem in nitride-based blue emitters. Li et al. in 2013 demonstrated the use of a p-GaN hole injection layer (HIL) as means to improve hole injection efficiency, which has been associated with efficiency droop.[88] The implementation of the p-GaN HIL reduced efficiency droop by 33% at 100 A/cm^2, ultimately improving output power and overall efficiency of the LED.[88] Given these promising results, HILs have been implemented with several materials, including p-GaN, p-InAlGaN, and p-AlGaN.

Although most blue LEDs have been fabricated using *c*-plane III-nitride materials, growth in the polar *c*-direction gives rise to polarization-induced electric fields, which cause a spatial separation of the electron and hole wave functions, and in turn a red shift of peak emission wavelengths due to the Stark effect, a blue shift of emission with increasing current density, and reduced efficiency because of the lessening of the wave function overlap for electrons and holes. These considerations have generated strong interest in the development of III-nitride LEDs on nonpolar orientations. To this end, Chakraborty et al.[89] have demonstrated InGaN/GaN LEDs grown by MOVPE on free-standing *m*-plane GaN. In their work, the free-standing *m*-plane template layer was grown by hydride vapor phase epitaxy (HVPE). The resulting InGaN/GaN quantum well device exhibited an emission peak at 450 nm with a continuous wave output of 250 μ*W* at a drive current of 20 mA.

9.5.5 Ultraviolet Light-Emitting Diodes

The ultraviolet portion of the spectrum ranges from 10 to 400 nm, and is further subdivided into UV-A (320–400 nm), UV-B (290–320 nm), UV-C (200–290 nm), and vacuum UV (10–200 nm). III-Nitride LEDS can cover the entire UV-A, UV-B, and UV-C range with the bandgaps accessible using InGaAlN. Figure 9.13 shows the typical structure of an InGaAlN-based ultraviolet-emitting diode. The substrate is usually sapphire, necessitating the use of a low-temperature buffer layer of AlN or GaN. The n-contact layer and n-confinement layers are both AlGaN. Though the p-confinement layer is also AlGaN, the p-contact layer is made from GaN to facilitate higher doping concentrations and lower specific resistivity. The active region, sandwiched between the p-type and n-type confinement layers, comprises undoped AlGaN barriers and InGaAlN quantum wells. For UV-A wavelengths longer than ~365 nm, the quantum wells can be InGaN, but for shorter wavelengths, it is necessary to include aluminum in the wells.

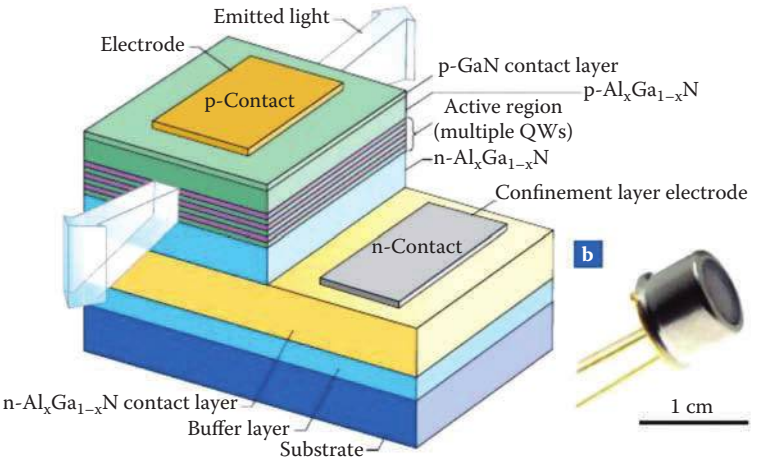


FIGURE 9.13
InGaAlN-based ultraviolet LED: (a) device structure and (b) packaged device. (Reprinted from A. Khan et al., *Nat. Photonics*, 2, 77 [2008]. With permission. Copyright 2008, Macmillan Publishers.)

The development of UV-A LEDs with an emission wavelength of >365 nm (corresponding to the band edge of GaN) is fairly mature and has followed the development of blue LEDs based on similar structures with InGaN MQWs. As with blue LEDs, these devices are grown on thick GaN layers deposited on lattice-mismatched sapphire (or sometimes SiC) substrates, and the resulting threading dislocation densities are very high (~10^8 to 10^9 cm^{-2}). The external quantum efficiency has a strong dependence on the threading dislocation density,[90,91] and this effect is more pronounced as the quantum well indium concentration is decreased to shorten the wavelength. This has resulted in a strong impetus for the development of high-quality GaN substrates with large area, and there have been some demonstrations of UV-A LEDs grown on GaN substrates with improved external quantum efficiency compared with their counterparts grown on sapphire.

UV-A LEDS with emission less than 365 nm present additional challenges because of the need to use aluminum-containing quantum wells. First, the quantum-confined Stark effect is more pronounced in aluminum-containing wells. The spontaneous and piezoelectric fields developed perpendicular to the (0001) surface tend to spatially separate the electrons and holes in the quantum wells, reducing their rate of radiative recombination. This can be alleviated to a great extent by using either semipolar or nonpolar crystal orientations.[92] Second, carrier confinement is poorer when higher-bandgap aluminum-containing wells are used. UV LEDs tend to show unwanted long-wavelength emission (365 nm) associated with recombination outside of the active layer. Suppression of this unwanted emission requires the insertion of an EBL between the active region and the n-GaN contact. Third, for UV-A LEDs with emission less than 365 nm, the thick GaN buffer and p-GaN contact layer both become absorbing, resulting in reduced extraction efficiency.

UV-B and UV-C LEDs are especially challenging because the high aluminum content precludes their growth on GaN buffer layers. This is because the difference in thermal expansion coefficients between GaN and AlN leads to high tensile stresses and cracking in AlGaN films with high aluminum content deposited on thick GaN buffers (or GaN substrates). Instead, it is necessary to grow thick AlN or AlGaN buffer layers on sapphire substrates. This has proven to be problematic because the surface mobility of aluminum adatoms is considerably less than that of gallium adatoms. This results in poor crystal

quality, and threading dislocation densities for AlN on sapphire (10^{10} to 10^{11} cm^{-2}) are two orders of magnitude higher than those for GaN on sapphire (10^{8} to 10^{9} cm^{-2}).[93] Several approaches have been attempted to achieve high-quality AlN-on-sapphire or SiC substrates. Pulsed atomic layer epitaxy (PALE), using pulsed modulation of the column III and nitrogen precursors, can enhance the surface migration of aluminum and gallium to improve crystal quality and morphology. PALE also avoids adduct formation by the reaction of NH_3 with the column III precursors trimethylgallium and trimethylaluminum, and gas phase adduct formation has been shown to degrade the surface morphology. High-temperature metalorganic vapor phase epitaxial growth with a low V-III ratio, growth on patterned substrates, and ELO have all been shown to lead to lower threading dislocation densities in AlN buffers. Balakrishnan et al.[94] reported growth of AlN on 6H-SiC with a threading dislocation density of only 10^{6} cm^{-2} using high-temperature MOVPE growth on patterned 6H-SiC substrates at 1500°C.

Because of the difficulties encountered with higher aluminum content devices, the measured external quantum efficiencies decrease with the wavelength of emission, as shown in Figure 9.14.

9.5.6 White Light-Emitting Diodes

White LEDs are of great interest for liquid crystal display (LCD) backlights, streetlights, automobile headlamps, large-area displays, and home and office illumination. Three figures of merit for white LEDS are the luminous efficacy—its ability to produce a visual sensation of brightness—in lumens; its color temperature, referenced to an ideal white source; and the color rendering index (CRI), a measure of the LED's ability to reproduce colors of objects as seen under the sun.

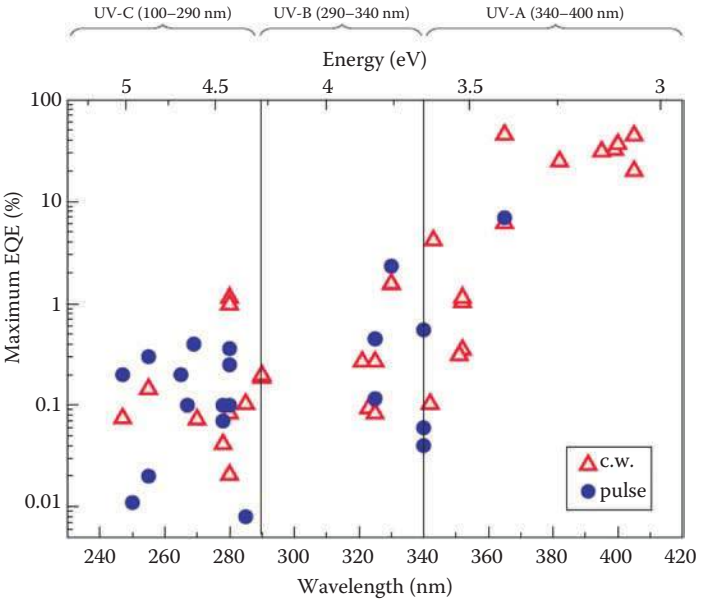


FIGURE 9.14
External quantum efficiencies (EQEs) for III-nitride-based UV LEDs, with continuous wave (c.w.) operation (open triangles) and pulsed operation (filled circles). (Reprinted from A. Khan et al., *Nat. Photonics*, 2, 77 [2008]. With permission. Copyright 2008, Macmillan Publishers.)

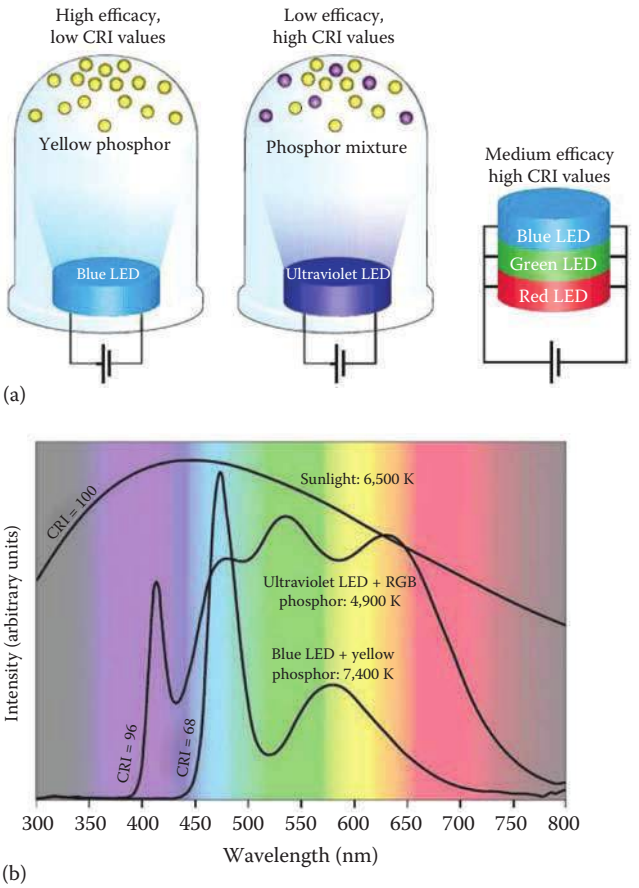


FIGURE 9.15
White LEDs. (a) Three common designs for white LEDs: a blue LED with a yellow phosphor, an ultraviolet LED with a phosphor mixture (red, green, and blue), and a combination of red, green, and blue LEDs. (b) Spectra of white LEDs compared with the solar spectrum. (Reprinted from S. Pimputkar et al., *Nat. Photonics*, 3, 180 [2009]. With permission. Copyright 2009, Macmillan Publishers.)

White LEDs are realized by one of three approaches,[95] shown in Figure 9.15. The first involves the mixing of blue and yellow light using a blue LED with a yellow phosphor. The second involves use of an ultraviolet LED with a mixture of red, green, and blue phosphors, and the third uses a combination of three separate LEDs emitting red, green, and blue.

Use of a blue LED with a yellow phosphor produces a cheap and efficient source of bright light, but the emitted spectrum is far from ideal. Here, the typical phosphor is $Y_3Al_5O_{12}:Ce^{3+}$, or cerium-doped yttrium-aluminum-garnet (YAG). All leaking light from the LED contributes to the visible spectrum, but the large amount of blue results in a high color temperature of ~7400 K. The high color temperature indicates a "cold" light, and the CRI is too low (~68) to be used for indoor lighting.

Multiphosphor devices may be made using ultraviolet or blue emitters. If an ultraviolet LED is combined with red, green, and blue phosphors, only downconverted light is in the visible region, so the efficiency and efficacy are relatively low. The spectrum is closer to ideal, with a color temperature of 4900 K ("warm" light) and a CRI of ~96. Use of a blue

LED with two phosphors, either green or yellow mixed with red, provides warm white light with a high CRI for indoor lighting applications. The efficacy can be higher than for the case of an ultraviolet emitter, because some of the blue emitted light is used directly without the need for downconversion.

The combination of red, green, and blue LEDs in a single package could produce a high CRI with good efficacy, but the challenge here is that green LEDs currently have poor efficiency compared with red and blue devices.

9.6 Solar Cells

Single-junction solar cells are capable of theoretical efficiencies of 30% under 1 sun illumination[96] and 37% at 1000 suns.[97] Metamorphic GaAs-on-Si(001) single-junction solar cells have been demonstrated with 7% efficiency under 1 sun,[98] but are limited by the threading dislocation density and associated nonradiative lifetime. The choice of bandgap involves a trade-off between thermalization losses and losses due to unabsorbed photons. Higher efficiencies can therefore be obtained using multiple energy gaps in tandem solar cells, and the maximum theoretical efficiency for a triple-junction solar cell is 56% at 1000 suns.[96] In a lattice-matched solar cell, the restriction of lattice matching dictates suboptimal combinations of bandgap,[99] whereas metamorphic realization lifts these restrictions and will enable higher efficiencies if low dislocation densities and low nonradiative recombination losses may be attained.[100]

State-of-the-art lattice-matched triple-junction solar cells involve the growth of two III-V junctions on top of a Ge substrate, which is used to make the bottom junction. Figure 9.16

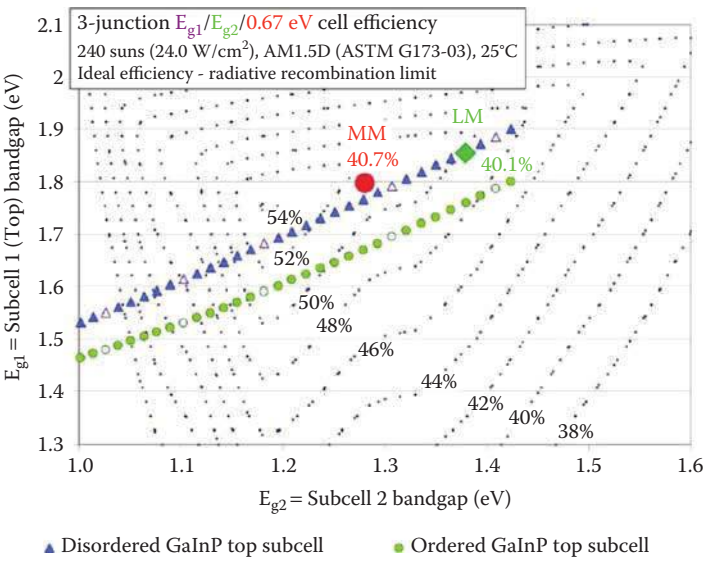


FIGURE 9.16
Isoefficiency contours calculated for triple-junction solar cells on Ge substrates as a function of the top cell bandgap (E_{g1}) and the middle cell bandgap (E_{g2}) with 240 suns illumination and the AM1.5D terrestrial solar spectrum. (Reprinted from R. R. King et al., *Appl. Phys. Lett.*, 90, 183516 [2007]. With permission. Copyright 2007, American Institute of Physics.)

shows the contours of the ideal triple-junction efficiency at 240 suns as a function of the top (E_{g1}) and middle junction (E_{g2}) bandgaps, with the assumption that the bottom junction is made in the Ge substrate with a bandgap of 0.67 eV. The bandgaps of the InGaP (E_{g1}) and InGaP (E_{g2}) are also plotted for the case of InGaP with an ordered column III sublattice (small filled triangles) and for the case of InGaP with a disordered group III sublattice (small filled circles), with the assumption that InGaAs and InGaP are lattice matched to one another. The lattice-matched triple-junction cell corresponds to the special case in which all three lattice constants are matched. Also shown in the figure are the champion lattice-matched and metamorphic cells on Ge substrates.

A maximum efficiency of 40.1% at 240 suns was obtained in a lattice-matched triple-junction cell employing a top $Ga_{0.5}In_{0.5}P$ junction and a middle $In_{0.02}Ga_{0.98}As$ junction, for a bandgap combination of 1.86, 1.39, and 0.67 eV, as shown by the large filled diamond in Figure 9.16.[99] A higher efficiency was obtained in a metamorphic triple-junction solar cell, in which the bandgaps of the top two cells may be varied without the restriction of lattice matching to the Ge substrate. In the champion metamorphic design shown by the large filled circle in Figure 9.16, increased indium content in the top two junctions reduced their bandgaps to 1.8 and 1.3 eV, resulting in a slightly higher efficiency of 40.7% at 240 suns.[99]

A shortcoming of the use of Ge substrates involves the nonoptimum bandgap of Ge(0.67 eV) for use in the bottom junction. Geisz et al. used a semiempirical approach to calculate isoefficiency surfaces for triple-junction solar cells assuming 500 suns and a perfect antireflective coating.[99] Figure 9.17 shows the results of these calculations, in which all three bandgaps were allowed to vary freely. The global maximum efficiency is expected with bandgaps of 1.86, 1.34, and 0.93 eV, while there is also a local maximum at 1.75, 1.18, and 0.70 eV. Either of these designs is expected to achieve about 5% higher efficiency than a lattice-matched design using a Ge substrate, and a key difference is that the bottom Ge cell exhibits higher thermalization losses due to its lower-than-optimal bandgap.

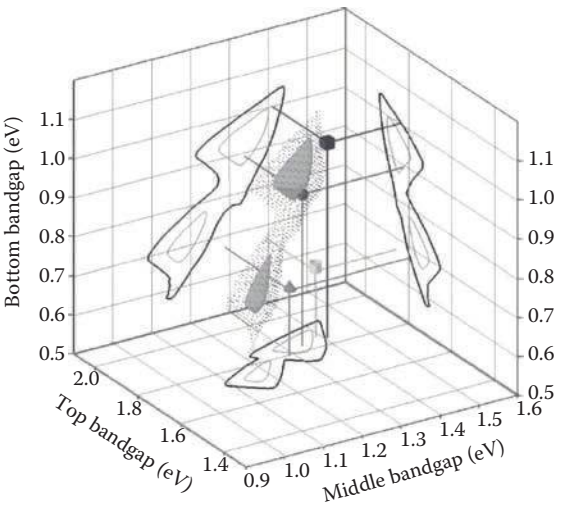


FIGURE 9.17
Theoretical isoefficiency surfaces for metamorphic triple-junction solar cells operated at 500 suns concentration. Champion devices that have been fabricated are shown by the filled cylinder (1.86, 1.39, and 1.00 eV), filled cone (1.80, 1.29, and 0.67 eV), filled cube (1.84, 1.41, and 1.00 eV), and filled sphere (1.85, 1.34, and 0.89 eV). (Reprinted from J. F. Geisz et al., *Appl. Phys. Lett.*, 93, 123505 [2008]. With permission. Copyright 2008, American Institute of Physics.)

One approach to the fabrication of a metamorphic triple-junction solar cell involves inverted growth on a GaAs substrate. GaAs exhibits a lattice constant close to that of Ge, so lattice-matched growth of the top and middle junctions could be achieved using compositions similar to those of a Ge substrate. The bottom cell may be made in InGaAs, which is not lattice matched but may take on the optimum bandgap of ~0.93 eV. However, because the bottom junction is grown last, the structure must be inverted and it is necessary to remove the GaAs substrate from the solar cell.

In such an inverted triple-junction cell, it is possible to lattice mismatch two or even all three junctions with respect to the GaAs substrate. This allows greater freedom in choosing the bandgaps of the three junctions, but this freedom must be balanced against the increased introduction of threading dislocations and the associated degradation in nonradiative minority carrier lifetime.

Geisz et al.[99] have demonstrated an inverted triple-junction solar cell grown on a GaAs substrate by MOVPE using two independently metamorphic junctions and achieved 40.8% efficiency at 326 suns. Figure 9.18 shows a focused-ion-beam cross-sectional image of this solar cell before removal of the GaAs substrate. The bottom junction is fabricated in InGaAs(37% indium, 0.89 eV), the middle junction is made in InGaAs(4% indium, 1.34 eV), and the top junction is fabricated in InGaP(50% indium, 1.83 eV). To accommodate the lattice mismatch, a graded AlInGaP buffer was inserted between the top and middle cells, and an InGaP graded buffer was inserted between the middle and bottom cells. Tunnel junctions were made at the interface between the top cell and graded AlInGaP buffer, and also between the middle cell and the InGaP graded buffer.

The design of such a cell is rather complex because the graded layers must not only accommodate the lattice mismatch but also should be transparent to the light passing

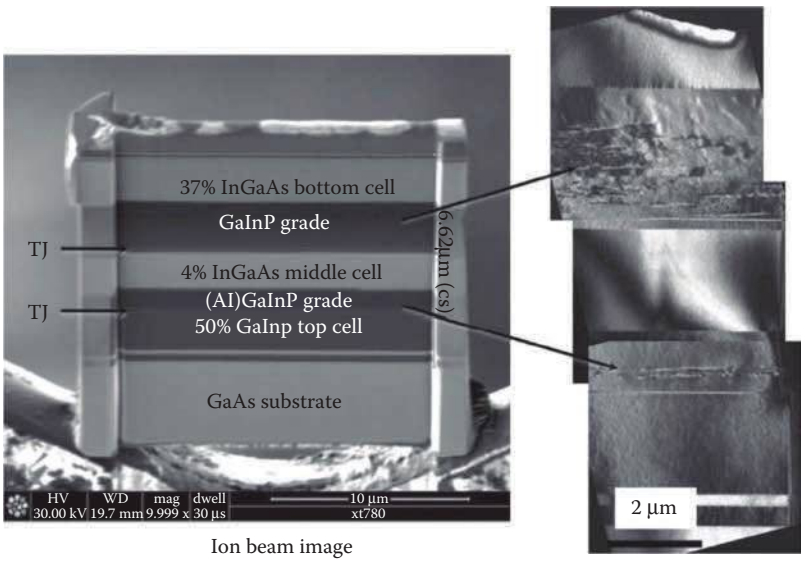


FIGURE 9.18
Inverted triple-junction metamorphic solar cell grown on a GaAs substrate using MOVPE. The left-hand side shows a focused-ion-beam cross-sectional image of the structure before removal from the GaAs substrate. The right-hand side shows corresponding cross-sectional TEM images. (Reprinted from J. F. Geisz et al., *Appl. Phys. Lett.*, 93, 123505 [2008]. With permission. Copyright 2008, American Institute of Physics.)

through them. Moreover, the graded layers should be capable of heavy doping so that the tunnel junctions operate at the high current densities occurring with high levels of concentration. In the solar cell fabricated by Geisz et al.[99] the tunnel junction between the top and middle junctions was lattice matched to the substrate and top junction, while the tunnel junction between the bottom and middle junctions was fabricated using $In_{0.04}Ga_{0.96}As$:Se/$Al_{0.2}Ga_{0.8}As$:C in which the $Al_{0.2}Ga_{0.8}As$:C layer was grown in tension. Despite this, the tunnel junction between the bottom and middle cells exhibited ohmic behavior up to 826 suns. Beyond this, the photocurrent exceeded the maximum tunneling current.

Structural analysis of a cell identical in structure to the one shown in Figure 9.18 was performed to determine the strains and dislocation densities in the material. X-ray diffraction analysis revealed that the metamorphic junctions contained 0.029% and 0.014% compressive strain, for the bottom and middle junctions, respectively. The threading dislocation densities, as estimated from spectrally resolved cathodoluminescence, were 10^5 cm^{-2} and 2–3 × 10^6 cm^{-2} in the middle and bottom junctions, respectively. The threading dislocation density in the top lattice-matched junction was reported to be below the detection limit for the cathodoluminescence technique.

Figures 9.19 and 9.20 show the measured performance of the inverted metamorphic triple-junction cell reported by Geisz et al.[99] The fill factor and efficiency fall off dramatically at greater than 841 suns because of the limitations of the bottom tunnel junction.

Metamorphic solar cells employing more than three junctions are also possible. Henry has reported on the ideal limiting efficiencies of multiple gap solar cells employing up to 36 junctions.[97] With a concentration of 1000 suns and a temperature of 300 K, the maximum efficiency is predicted to be 37%, 50%, 56%, and 72% using 1, 2, 3, and 36 bandgaps, respectively. In practice, however, it may be increasingly difficult to approach the theoretical efficiencies with a larger number of bandgaps because of the need for (N − 1) tunnel junctions in a series-connected cell involving N bandgaps.

Stan et al.[101] demonstrated an inverted quad-junction solar cell grown on a GaAs substrate by MOVPE with two metamorphic junctions. The structure of this cell is shown schematically in Figure 9.21. The four junctions are made using InGaP, GaAs, InGaAs, and InGaAs with bandgaps of 1.91, 1.42, 1.02, and 0.7 eV, respectively. The InGaP and GaAs

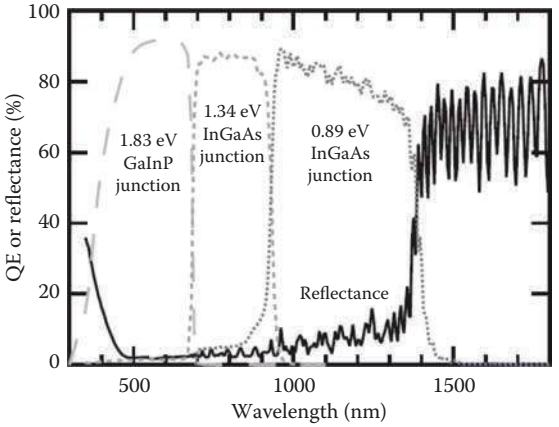


FIGURE 9.19
Quantum efficiency (QE) and reflectance of an inverted metamorphic triple-junction solar cell fabricated by Geisz et al. (Reprinted from J. F. Geisz et al., *Appl. Phys. Lett.*, 93, 123505 [2008]. With permission. Copyright 2008, American Institute of Physics.)

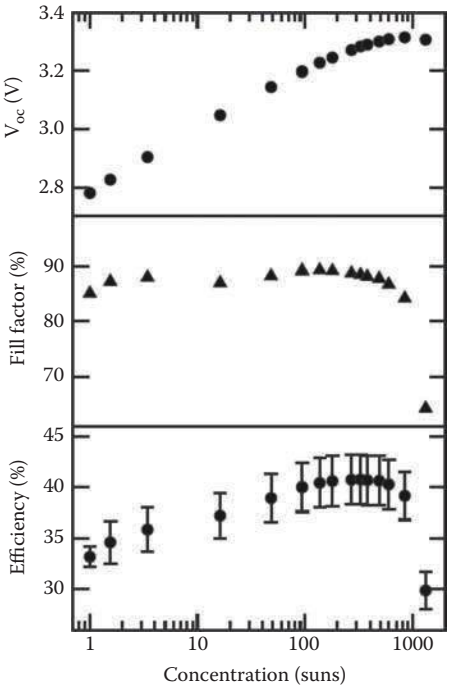


FIGURE 9.20
Efficiency, fill factor, and open circuit voltage of an inverted metamorphic triple-junction solar cell fabricated by Geisz et al. (Reprinted from J. F. Geisz et al., *Appl. Phys. Lett.*, 93, 123505 [2008]. With permission. Copyright 2008, American Institute of Physics.)

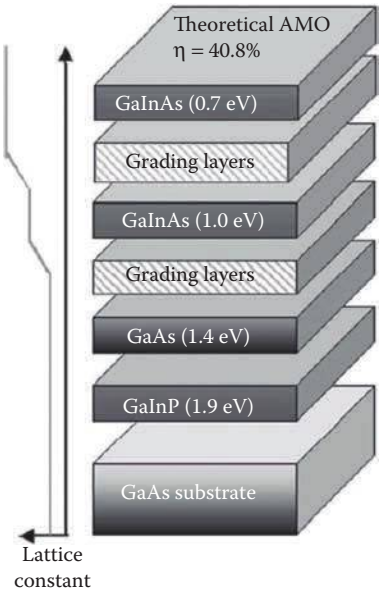


FIGURE 9.21
Structure of quad-junction inverted metamorphic solar cell fabricated by Stan et al. (Reprinted from M. Stan et al., *J. Cryst. Growth*, 312, 1370 [2010]. With permission. Copyright 2010, Elsevier.)

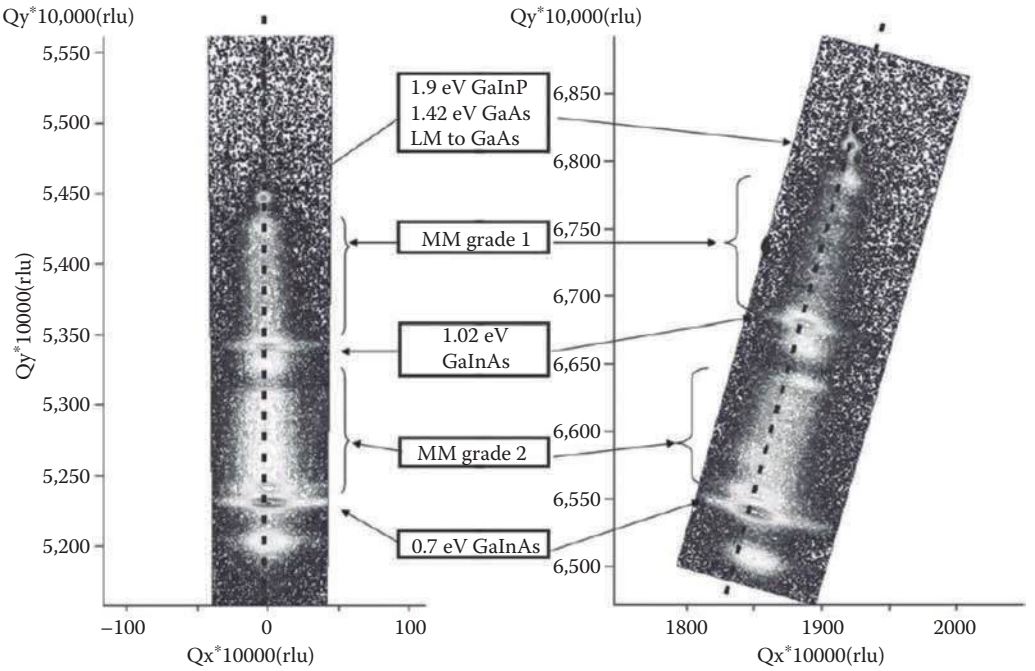


FIGURE 9.22
High-resolution 004 and 115 x-ray reciprocal space maps for an inverted quad-junction metamorphic solar cell demonstrated by Stan et al. (Reprinted from M. Stan et al., *J. Cryst. Growth*, 312, 1370 [2010}. With permission. Copyright 2010, Elsevier.)

junctions are lattice matched to the substrate and require no graded buffers for their growth, but the lattice-mismatched InGaAs junctions are realized using graded InGaAs. This cell achieved a record 1 sun AM0 conversion efficiency of 33.6%, and the predicted AM1.5D 500 sun efficiency is 43.0%.[102]

Figure 9.22 shows the high-resolution x-ray reciprocal space maps of the inverted quad-junction metamorphic solar cell demonstrated by Stan et al.,[100] obtained using a Philips MRD diffractometer and Cu $k\alpha_1$ radiation. The dashed lines shown on the reciprocal space maps indicate the conditions of tilt-free lattice-relaxed epitaxy. The observed departures from the dashed lines indicate tensile strain of 5.25×10^{-4} in the 1.02 eV subcell and compressive strain of 5.56×10^{-4} in the 0.7 eV subcell.

Figure 9.23 shows the wavelength-resolved cathodoluminescence images of the same inverted quad-junction solar cell.[100] Figure 9.23a–c corresponds to the 1.0 eV InGaAs junction, the 1.42 eV GaAs junction, and the 1.91 eV InGaP junction, respectively. The threading dislocation densities estimated from these images are 5×10^6 cm^{-2} for the lattice-mismatched 1.0 eV InGaAs subcell, but 5×10^5 cm^{-2} for the lattice-matched 1.42 eV GaAs cell.

Figure 9.24 shows the external quantum efficiency and reflectance characteristics for the quad-junction metamorphic solar cell.[100] The solar cell area was 4 cm^2 and AM0 illumination was used. The results indicate that the 0.7 eV cell is current limiting, and that further optimization is possible.

Other metamorphic multijunction solar cell designs are possible and could have high theoretical efficiencies. For example, conventional (noninverted) triple-junction or quad-junction structures could be fabricated using an InGaAs bottom junction grown directly

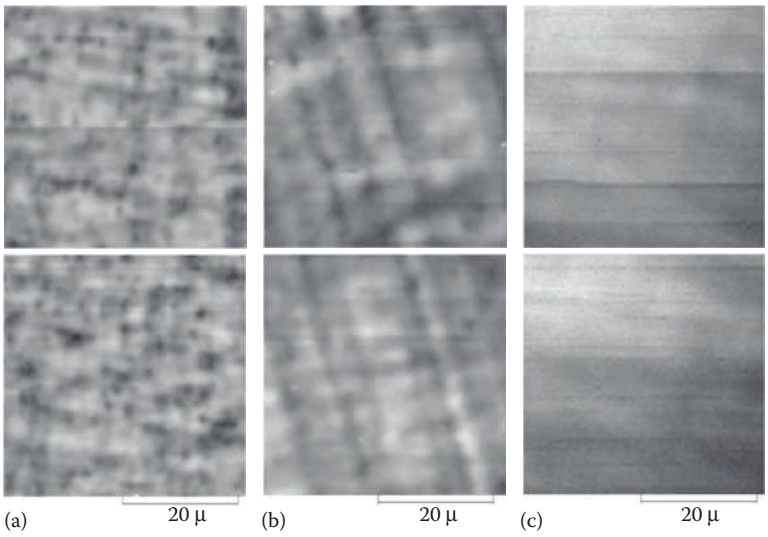


FIGURE 9.23
Cathodoluminescence images from an inverted quad-junction metamorphic solar cell demonstrated by Stan et al.: (a) 1.0 eV InGaAs cell, (b) 1.42 eV GaAs cell, and (c) 1.91 eV InGaP cell. (Reprinted from M. Stan et al., *J. Cryst. Growth*, 312, 1370 [2010]. With permission. Copyright 2010, Elsevier.)

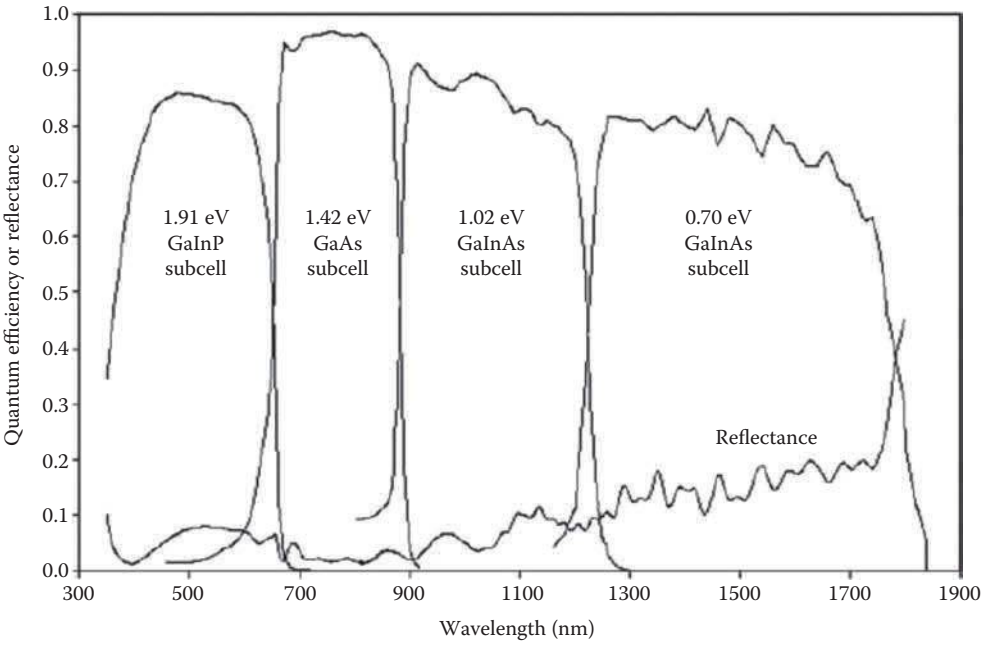


FIGURE 9.24
External quantum efficiency and reflectance of an inverted quad-junction metamorphic solar cell demonstrated by Stan et al.[100] (Reprinted from M. Stan et al., *J. Cryst. Growth*, 312, 1370 [2010]. With permission. Copyright 2010, Elsevier.)

on the GaAs substrate. Use of four to six junctions might enable higher efficiencies than those obtained in triple-junction cells, but requires a wide bandgap (2.0–2.2 eV) for the top junction. Unfortunately, most materials with the necessary bandgap are indirect gap or contain aluminum, which results in low efficiency due to the tendency to form oxygen complex deep levels. InGaP may be used as the wide-bandgap material without the aforementioned disadvantages, but is lattice mismatched to GaAs.[103] Although the noninverted structures and cells with more than three junctions involve all mismatched materials, the use of appropriate graded buffers could make such designs practical.

9.7 Conclusions

This chapter, though far from exhaustive, illustrates that metamorphic growth has been applied extensively to every important class of semiconductor device. The advantage of flexibility in device design often outweighs the need for lattice matching. Thus, devices with excellent performance have been developed—and in many cases commercialized—using metamorphic growth on mismatched substrates, despite issues of strain and lattice defects. It is likely that this trend will continue, and that graded layers and defect engineering approaches will make it possible to solve the important problems associated with mismatched growth.

References

1. A. G. O'Neill and D. A. Antoniadis, Deep submicron cmOS based on silicon germanium technology, *IEEE Trans. Electron. Dev.*, 43, 911 (1996).
2. M. V. Fischetti and S. E. Laux, Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys, *J. Appl. Phys.*, 80, 2234 (1996).
3. F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, Mechanism and conditions for anomalous strain relaxation in graded thin films and superlattices, *J. Appl. Phys.*, 71, 4230 (1992).
4. M. Behet, K. van der Zanden, G. Borghs, and A. Behres, Metamorphic InGaAs/InAlAs quantum well structures grown on GaAs substrates for high electron mobility transistor applications, *Appl. Phys. Lett.*, 73, 2760 (1998).
5. D. E. Grider, S. E. Swirhun, D. H. Narum, A. I. Akinwande, T. E. Nohava, W. R. Stuart, P. Joslyn, and K. C. Hsieh, Metamorphic $In_yGa_{1-y}As/In_zAl_{1-z}As$ heterostructures field effect transistors grown on GaAs(001) substrates using molecular-beam epitaxy, *J. Vac. Sci. Technol. B*, 8, 301 (1990).
6. T. Mishima, K. Higuchi, M. Mori, and M. Kudo, High G_m $In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ high electron mobility transistors grown lattice-mismatched on GaAs substrates, *J. Cryst. Growth*, 150, 1230 (1995).
7. S. Gozu, K. Tsuboki, M. Hayashi, C. Hong, and S. Yamada, Very high electron mobilities at low temperatures in $In_xGa_{1-x}As/In_yAl_{1-y}As$ HEMTs grown lattice-mismatched on GaAs substrates, *J. Cryst. Growth*, 201–202, 749 (1999).
8. L. J. Cui, Y. P. Zeng, B. Q. Wang, J. Wu, Z. P. Zhu, and L. Y. Lin, Rapid thermal annealing effects on step-graded InAlAs buffer layer and $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ metamorphic high electron mobility transistor structures on GaAs substrates, *J. Appl. Phys.*, 91, 2429 (2002).

9. Ch. Heyn, S. Mendach, S. Lohr, S. Beyeer, S. Scnull, and W. Hansen. Growth of shallow InAs HEMTs with metamorphic buffer, *J. Cryst. Growth*, 251, 832 (2003).
10. M. Haupt, K. Kohler, P. Ganser, S. Emminger, S. Muller, and W. Rothemund, Growth of high-quality $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures using strain relaxed $Al_xGa_yIn_{1-x-y}As$ buffer layers on GaAs, *Appl. Phys. Lett.*, 69, 412 (1996).
11. K. Inoue, J. C. Harmand, and T. Matsuno, High-quality $In_xGa_{1-x}As$/InAlAs modulation-doped heterostructures grown lattice-mismatched on GaAs substrates, *J. Cryst. Growth*, 111, 313 (1991).
12. S.-K. Hong, H.-G. Lee, J.-J. Lee, S.-G. Kim, K.-E. Pyun, and H.-M. Park, Molecular beam epitaxy growth of indium-rich $In_xGa_{1-x}As/In_yAl_{1-y}As$/InP structures towards high channel conductivity for a high electron mobility transistor using a linearly graded buffer layer, *J. Cryst. Growth*, 169, 435 (1996).
13. M. Haupt, K. Kohler, P. Ganser, S. Muller, and W. Rothemund, Molecular beam epitaxy of $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ heterostructures on metamorphic $Al_xGa_yIn_{1-x-y}As$ buffer layers, *J. Cryst. Growth*, 175–176, 1028 (1997).
14. A. S. Wakita, H. Rohdin, V. M. Robbins, N. Moll, C.-Y. Su, A. Nagy, and D. P. Basile, Low-noise bias reliability of AlInAs/GaInAs MODFETs with linearly graded low-temperature buffer layers grown on GaAs substrates, presented at 10th International Conference on Indium Phosphide and Related Materials, Tsukuba, Japan (May 1998).
15. A. Wakita, H. Rohdin, V. Robbins, N. Moll, C.-Y. Su, A. Nagy, and D. Basile, Low-noise bias reliability of AlInAs/GaInAs modulation-doped field effect transistors with linearly graded low-temperature buffer layers grown on GaAs substrates, *Ipn. J. Appl. Phys.*, 38, 1186 (1999).
16. Y. Cordier, J.-M. Chauveau, D. Ferre, and J. Dipersio, Comparison of $In_{0.33}Al_{0.67}As/In_{0.34}Ga_{0.66}AS$ on GaAs metamorphic high electron mobility transistors grown by molecular beam epitaxy wih normal and inverse step on linear graded buffer layers, *J. Vac. Sci. Technol. B*, 18, 2513 (2000).
17. M. Dammann, M. Chertouk, W. Jantz, K. Kohler, and G. Weimann, Reliability of InAlAs/InGaAS HEMTs grown on GaAs substrate with metamorphic buffer, *Microelectron. Reliab.*, 40, 1709 (2000).
18. K. Yuan, K. Radhakrishnan, H. Q. Zheng, and G. I. Ng, Metamorphic $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ high electron mobility transistors on GaAs with $In_xGa_{1-x}P$ graded buffer, *J. Vac. Sci. Technol. B*, 19, 2119 (2001).
19. W. E. Hoke, T. D. Kennedy, A. Torabi, C. S. Whelan, P. F. Marsh, R. E. Leoni, S. M. Lardizabal, et al., Properties of metamorphic materials and device structures on GaAS substrates, *J. Cryst. Growth*, 251, 804 (2003).
20. E. S. Semenova, A. E. Zhukov, S. S. Mikhrin, A. Yu. Egorov, V. A. Odnoblyudov, A. P. Vail'ev, E. V. Nikitina, et al., Metamorphic growth for supplication in long-wavelength (1.3–1.55 um) lasers and MODFET-type structures on GaAs substrates, *Nanotechnology*, 15, S283 (2004).
21. H. Choi, Y. Jeong, J. Cho, and M. H. Jeon, Effectiveness of non-linear graded buffers for In(Ga,Al)AS metamorphic layers grown on GaAs(001), *J. Cryst. Growth*, 311, 1091 (2009).
22. J.-H. Tsai, Study of metamorphic co-integrated heterostructures bipolar and field effect transistors (BIFETs), *J. Electrochem. Soc.*, 158, H889 (2011).
23. G. B. Galiev, I. S. Vasil'evskii, S. S. Pushkarev, E. A. Klimov, R. M. Imamov, P. A. Buffat, B. Dwir, and E. I. Suvorova, Metamorphic InAlAs/InGaAs/InAl/As/GaAs HEMT heterostructures containing strained superlattices and inverse steps in the metamorphic buffer, *J. Cryst. Growth*, 366, 55 (2013).
24. S. Mendach, C. M. Hu, Ch. Heyn, S. Schnull, H. P. Oepen, R. Anton, and W. Hansen, Strain relaxation in high-mobility InAs inserted-channel heterostructures with metamorphic buffer, *Physica E*, 13, 1204 (2002).
25. H.-R. Blank, M. Thomas, K. C. Wong, and H. Kroemer, Influence of the buffer layers on the morphology and transport properties in InAl/(Al,Ga)Sb quantum wells grown by molecular beam epitaxy, *Appl. Phys. Lett.*, 69, 2080 (1996).
26. M. Thomas, H.-R. Blank, K. C. Wong, and H. Kroemer, Buffer-dependent mobility and morphology of InAs/(Al,Ga)Sb quantum wells, *J. Cryst. Growth*, 175/176, 894 (1997).

27. C.-S. Lee and W.-T. Chien, Investigations on InP/InAlAsSb/GaAs metamorphic high electron mobility transistors with a dual-composition-graded $In_xGa_{1-x}As_{1-y}Sb_y$ channel and different Schottky-barrier gate structures, *J. Electrochem. Soc.*, 158, H452 (2011).
28. Y. Song, S. Wong, I. Tangring, Z. Lai, and M. Sadeghi, Effects of doping and grading slope on surface and structure of metamorphic InGaAs buffers on GaAs substrates, *J. Appl. Phys.*, 106, 123531 (2009).
29. J.-I. Chyi, J.-L. Shieh, J.-W. Pan, and R.-M. Lin, Material properties of compositional graded $In_xGa_{1-x}As$ and $In_xAl_{1-x}As$ epilayers grown on GaAs substrates, *J. Appl. Phys.*, 79, 8367 (1996).
30. S. T. Sheppard, K. Doverspike, W. L. Pribble, S. T. Allen, J. W. Palmour, L. T. Kehias, and T. J. Jenkins, High-power microwave GaN/AlGaN HEMT's on semi-insulating silicon carbide substrates, *IEEE Electron Dev. Lett.*, 20, 161 (1999).
31. Y.-F. Wu, D. Kapolnek, J. P. Ibbetson, P. Parikh, B. P. Keller, and U. K. Mishra, Very-high power density AlGaN/GaN HEMTs, *IEEE Trans. Electron Dev.*, 48, 586 (2001).
32. T. Kikkawa, T. Maniwa, H. Hayashi, M. Kanamura, S. Yokokawa, M. Nishi, N. Adachi, M. Yokoyama, Y. Tateno, and K. Joshin, An over 200-W output power GaN HEMT push-pull amplifier with high reliability, in *Microwave Symposium Digest, 2004 IEEE MTT-S International Symposium* (June 6–11, 2004), p. 1347.
33. L. Shen, R. Coffie, D. Buttari, S. Heikman, A. Chakraborty, A. Chini, S. Keller, S. P. DenBaars, and U. K. Mishra, High-power polarization-engineered GaN/AlGaN/GaN HEMTs without surface passivation, *IEEE Electron Dev. Lett.*, 25, 7 (2004).
34. A. Edwards, J. A. Mittereder, S. C. Binari, D. S. Katzer, D. F. Storm, and J. A. Roussos, Improved reliability of AlGaN-GaN HEMTs using an NH_3 plasma treatment prior to SiN passivation, *IEEE Electron Dev. Lett.*, 26, 225 (2005).
35. T. Palacios, A. Chakraborty, S. Heikamn, S. Keller, S. P. DenBaars, and U. K. Mishra, AlGaN/GaN high electron mobility transistors with InGaN back-barriers, *IEEE Electron Dev. Lett.*, 27, 13 (2006).
36. E. Zanoni, F. Danesin, M. Meneghini, A. Cetronio, C. Lanzieri, M. Peroni, and G. Meneghesso, Localized damage in AlGaN/GaN HEMTs induced by reverse-bias testing, *IEEE Electron Dev. Lett.*, 30, 427 (2009).
37. J. W. Chung, W. E. Hoke, E. M. Chumbes, and T. Palacios, AlGaN/GaN HEMT with 300-GHz f_{max}, *IEEE Electron Dev. Lett.*, 31, 195 (2010).
38. A. Crespo, M. M. Bellot, K. D. Chabak, J. K. Gillespie, G. H. Jensen, V. Miller, M. Trejo, et al., High-power Ka-band performance of AlInN/GaN HEMT with 9.8-nm-thin barrier, *IEEE Electron Dev. Lett.*, 31, 2 (2010).
39. K. Shinohara, A. Corrion, D. Regan, I. Milosavljevic, D. Brown, S. Burnham, P. J. Willadsen, et al., 220 GHz f_T and 400 GHz f_{max} in 40-nm GaN DH-HEMTs with re-grown ohmic, *IEDM Tech. Dig.*, 30.1.1 (2010).
40. H. Sun, A. R. Alt, H. Benedickter, E. Feltin, J.-F. Carlin, M. Gonschorek, N. Grandjean, and C. R. Bolognesi, 205-GHz (Al,In)N/GaN HEMTs, *IEEE Electron Dev. Lett.*, 31, 957 (2010).
41. D. S. Lee, X. Gao, S. Guo, D. Kopp, P. Fay, and T. Alacios, 300-GHz InAlN/GaN HEMTs with InGaN back barrier, *IEEE Electron Dev. Lett.*, 32, 1525 (2011).
42. A. Malmos, P. Gamarra, M.-A. di Forte-Poisson, H. Hjelmgren, C. Lacam, M. Thorsell, M. Tordjman, R. Aubry, and N. Rorsman, Evaluation of thermal versus plasma-assisted ALD Al_2O_3 as passivation for InAlN/AlN/GaN HEMTs, *IEEE Electron Dev. Lett.*, 36, 235 (2015).
43. B. M. Green, K. K. Chu, E. M. Chumbes, J. A. Smart, J. R. Shealy, L. F. Eastman, The effect of surface passivation on the microwave characteristics of undoped AlGaN/GaN HEMT's, *IEEE Electron Dev. Lett.*, 21, 268 (2000).
44. S. Binari, K. Ikossi, J. A. Roussos, W. Kruppa, D. Park, H. B. Dietrich, D. D. Koleske, A. Wickendon, and R. L. Henry, Trapping effects and microwave power performance in AlGaN/GaN HEMTs, *IEEE Trans. Electron Dev.*, 48, 465 (2001).
45. N. Okamoto, K. Hoshino, N. Hara, M. Takikawa, and Y. Arakawa, MOCVD-grown InGaN-channel HEMT structures with electron mobility of over 1000 cm^2/Vs, *J. Cryst. Growth*, 272, 278 (2004).

46. Y. Cai, Y. Zhou, K. J. Chen, and K. M. Lau, High-performance enhancement-mode AlGaN/GaN HEMTs using fluoride-based plasma treatment, *IEEE Electron Dev. Lett.*, 26, 435 (2005).
47. J. Liu, Y. Zhou, J. Zhu, and K. M. Lau, AlGaN/GaN/InGaN/GaN DH-HEMTs with an InGaN notch for enhanced carrier confinement, *IEEE Electron Dev. Lett.*, 27, 10 (2006).
48. T. Mizutani, M. Ito, S. Kishimoto, and F. Nakamura, AlGaN/GaN HEMTs with thin InGaN cap layer for normally off operation, *IEEE Electron Dev. Lett.*, 28, 549 (2007).
49. W. Saito, T. Nitta, Y. Kakiuchi, Y. Saito, K. Tsuda, I. Omura, and M. Yamaguchi, A 120-W boost converter operation using a high-voltage GaN-HEMT, *IEEE Electron Dev. Lett.*, 29, 8 (2008).
50. D. C. Dumka, C. Lee, H. Q. Tserng, P. Saunier, and M. Kumar, AlGaN/GaN HEMTs on Si substrate with 7W/mm output power density at 10 GHz, *Electron. Lett.*, 40, 1023 (2004).
51. A. Minko, V. Hoel, E. Morvan, B. Grimbert, A. Soltani, E. Delos, D. Ducatteau, et al., AlGaN-GaN HEMTs on Si with power density performance of 1.9 W/mm at 10 GHz, *IEEE Electron Dev. Lett.*, 25, 453 (2004).
52. D. Ducatteau, A. Minko, V. Hoel, E. Morvan, E. Delos, B. Grimbert, H. Lahreche, P. Bove, C. Gaguiere, J. C. DeJaeger, and S. Delage, Output power density of 5.1/mm at 18GHz with an AlGaN/GaN HEMT on Si substrate, *IEEE Electron Dev. Lett.*, 27, 7 (2006).
53. S. L. Selvaraj, T. Suzue, and T. Egawa, Breakdown enhancement of AlGaN/GaN HEMTs on 4-in silicon by improving the GaN quality on thick buffer layers, *IEEE Electron Dev. Lett.*, 30, 587 (2009).
54. S. Huang, Q. Jiang, S. Yang, C. Zhou, and K. J. Chen, Effective passivation of AlGaN/GaN HEMTs by ALD-grown AlN thin film, *IEEE Electron Dev. Lett.*, 33, 516 (2012).
55. K. K. Chu, P. C. Chao, M. T. Pizzella, R. Actis, D. E. Meharry, K. B. Nichols, R. P. Vaudo, X. Xu, J. S. Flynn, J. Dion, and G. R. Brandes, 9.4-W/mm power density AlGaN-GaN HEMTs on free-standing GaN substrates, *IEEE Electron Dev. Lett.*, 25, 596 (2004).
56. P. M. Asbeck, E. T. Yu, S. S. Lau, G. J. Sullivan, J. Van Hove, and J. Redwing, Piezoelectric charge densities in AlGaN/GaN HFETs, *Electron. Lett.*, 33, 1230 (1997).
57. H. Yang, H. Wang, K. Radhakrishnan, and C. L. Tan, Thermal resistance of metamorphic InP-based HBTs on GaAs substrates using a linearly graded $In_xGa_{1-x}P$ metamorphic buffer, *IEEE Trans. Electron Dev.*, 51, 1221 (2004).
58. K. L. Lew, S. F. Yoon, H. Tanoto, K. P. Chen, C. L. Dohrman, D. M. Isaacson, and E. A. Fitzgerald, InGaP/GaAs heterojunction bipolar transistor grown on Si substrate with SiGe graded buffer layer, *Electron. Lett.*, 44, 243 (2008).
59. A. Cavus, C. Monier, C. Cox, D. Pascua, A. Gutierrez-Aitken, A. Noori, S. Hayashi, and M. Goorsky, Development of 6.00 Å graded metamorphic buffer layers and high performance $In_{0.86}Al_{0.14}As/In_{0.86}Ga_{0.14}As$ heterojunction bipolar transistor devices, *J. Vac. Sci. Technol. B*, 24, 1492 (2006).
60. E. S. Yang, *Fundamentals of Semiconductor Devices* (New York: McGraw-Hill, 1978), p. 163.
61. M. R. Krames, O. R. Shchkekin, R. Mueller-Mach, G. O. Mueller, L. Zhou, G. Harbers, and M. G. Craford, Status and future of high-power light-emitting diodes for solid state lighting, *J. Display Technol.*, 3, 160 (2007).
62. S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. (Hoboken, NJ: John Wiley & Sons, 2007), p. 614.
63. M. G. Craford and W. O. Groves, Vapor phase epitaxial materials for LED applications, *Proc. IEEE*, 61, 862 (1973).
64. T. Sato and M. Imai, Characteristics of nitrogen-doped GaAsP light-emitting diodes, *Jpn. J. Appl. Phys.*, 41, 5995 (2002).
65. S. Kishino, M. Ogrima, and K. Kurata, A cross-hatch pattern in $GaAs_{1-x}P_x$ epitaxially grown on GaAs substrate, *J. Electrochem. Soc.*, 119, 617 (1972).
66. M. S. Abrahams, L. R. Weisberg, C. J. Buiocchi, and J. Blanc, Dislocation morphology in graded heterojunctions: $GaAs_{1-x}P_x$, *J. Mater. Sci.*, 4, 223 (1969).
67. D. A. Grenning and A. H. Herzog, Dislocations and their relation to irregularities in zinc-diffused GaAsP p-n junctions, *J. Appl. Phys.*, 39, 2783 (1968).
68. G. B. Stringfellow and P. E. Greene, Dislocations in $GaAs_{1-x}P_x$, *J. Appl. Phys.*, 40, 502 (1969).

69. B. W. Hakki, A. Jayaraman, and C. K. Kim, Band structure of InGaP from pressure experiments, *J. Appl. Phys.*, 41, 5291 (1970).
70. S. Kondo, H. Nagai, Y. Itoh, and M. Yamaguchi, InGaP orange light-emitting diodes on Si substrates, *Appl. Phys. Lett.*, 55, 1981 (1989).
71. M. Hoshino, K. Kitahara, K. Kodama, and M. Ozeki, GaInP growth by chloride vapor phase epitaxy, *J. Cryst. Growth*, 96, 188 (1989).
72. L. J. Stinson, J. G. Yu, S. D. Lester, M. J. Peanasky, and K. Park, High-efficiency InGaP light-emitting diodes on GaP substrates, *Appl. Phys. Lett.*, 58, 2012 (1991).
73. J. Nishizawa, M. Koike, and C. C. Jin, Efficiency of GaAlAs heterostructures red light-emitting diodes, *J. Appl. Phys.*, 54, 2807 (1983).
74. S. Giugni and T. L. Tansley, Comment on the compositional dependence of bandgap in AlGaAs and band-edge discontinuities in AlGaAs-GaAs heterostructures, *Semicond. Sci. Technol.*, 7, 1113 (1992).
75. O. Kwon, J. Boeckl, M. L. Lee, A. J. Pitera, E. A. Fitzgerald, and S. A. Ringel, Growth and properties of AlInGaP resonant cavity light emitting diodes on Ge/SiGe/Si substrates, *J. Appl. Phys.*, 97, 034504 (2005).
76. H. Sugawara, K. Itaya, H. Nozaki, and G. Hatakoshi, High-brightness InGaAlP green light-emitting diodes, *Appl. Phys. Lett.*, 61, 1775 (1992).
77. R. A. Logan, H. G. White, and W. Wiegmann, Efficient green electroluminescence in nitrogen-doped GaP p-n junctions, *Appl. Phys. Lett.*, 13, 139 (1968).
78. D. B. Eason, Z. Yu, W. C. Hughes, W. H. Roland, C. Boney, J. W. Cook Jr., and J. F. Schetzina, High-brightness blue and green light-emitting diodes, *Appl. Phys. Lett.*, 66, 115 (1995).
79. M. Hovinen, J. Ding, A. V. Nurmikko, G. C. Hua, D. C. Grillo, L. He, J. Han, and R. L. Gunshor, Degradation of (Zn,Cd)Se quantum well heterostructures for blue/green light emitters under optical injection, *Appl. Phys. Lett.*, 66, 2013 (1995).
80. Y. Zhu, S. McKerman, J. Xie, T. J. Miller, M. A. Haase, X. Sun, T. L. Smith, and C. A. Leatherdale, Effects of strain on defect structure in III-VI green color converters, *J. Appl. Phys.*, 108, 123104 (2010).
81. S. Nakamura, M. Senoh, N. Iwasa, S.-I. Nagahama, T. Yamada, and T. Mukai, Superbright green InGaN single-quantum-well-structure light-emitting diodes, *Jpn. J. Appl. Phys.*, 34, L1332 (1995).
82. D.-S. Oh, J.-J. Jang, O. Nam, K.-M. Song, and S.-N. Lee, Study of green light-emitting diodes grown on semipolar (11-22) GaN/m-sapphire with different crystal qualities, *J. Cryst. Growth*, 326, 33 (2011).
83. J. Kang, H. Li, Z. Li, Z. Liu, P. Ma, X. Yi, and G. Wang, Enhancing the performance of green GaN-based light-emitting diodes with graded superlattice AlGaN/GaN inserting layer, *Appl. Phys. Lett.*, 103, 102104 (2013).
84. H. Li, P. Li, J. Kang, Z. Li, Y. Zhang, M. Liang, Z. Li, J. Li, X. Yi, and G. Wang, Analysis model for efficiency droop of InGaN light-emitting diodes based on reduced effective volume of active region by carrier localization, *App. Phys. Exp.*, 6, 092101 (2013).
85. S. Nakamura and M. R. Krames, History of gallium-nitride-based light-emitting diodes for illumination, *Proc. IEEE*, 101, 10 (2013).
86. H. Amano, N. Sawaki, I. Akasaki, and Y. Toyoda, Metalorganic vapor phase epitaxial growth of a high-quality GaN film using an AlN buffer layer, *Appl. Phys. Lett.*, 48, 353 (1986).
87. S. Nakamura, T. Mukai, M. Senoh, and N. Iwasa, Thermal annealing effects on p-type Mg-doped GaN films, *Jpn. J. Appl. Phys.*, 31, L139 (1992).
88. H. Li, J. Kang, P. Li, J. Ma, H. Wang, M. Liang, Z. Li, J. Li, X. Yi, and G. Wang, Enhanced performance of GaN based light-emitting diodes with a low temperature p-GaN hole injection layer, *Appl. Phys. Lett.*, 102, 011105 (2013).
89. A. Chakraborty, B. A. Haskell, S. Keller, J. S. Speck, S. P. Denbaars, S. Nakamura, and U. K. Mishra, Demonstration of nonpolar m-plane InGaN/GaN light-emitting diodes on free-standing m-plane GaN substrates, *Jpn. J. Appl. Phys.*, 44, L173 (2005).

90. T. Makaui, S. Nagahama, N. Iwasa, M. Senoh, and T. Yamada, Nitride light emitting diodes, *J. Phys. Condens. Matter*, 13, 7089 (2001).
91. K. Akita, T. Kyono, Y. Yoshizumi, H. Kitabayashi, and K. Katayama, Characteristics of InGaN light-emitting diodes on GaN substrates with low threading dislocation densities, *Phys. Status Solidi A*, 204, 246 (2007).
92. A. Chakraborty, B. A. Haskell, S. Keller, J. S. Speck, S. P. DenBaars, S. Nakamura, and U. K. Mishra, Demonstration of nonpolar m-plane InGaN/GaN light-emitting diodes on free-standing m-plane GaN substrates, *Jpn. J. Appl. Phys.*, 44, L173 (2005).
93. M. Imura, K. Nakano, N. Fujimoto, N. Okada, K. Balakrishnan, M. Iwaya, S. Kamiyama, H. Amano, I. Akasaki, and T. Noro, Dislocations in AlN epilayers grown on sapphire substrate by high-temperature metal-organic vapor phase epitaxy, *Jpn. J. Appl. Phys.*, 46, 1458 (2007).
94. K. Balakrishnan, A. Bandoh, M. Iwaya, S. Kamiyama, H. Amano, and I. Akasaki, Influence of high temperature in the growth of low dislocation content AlN bridge layers on patterned 6H-SiC substrates by metalorganic vapor phase epitaxy, *Jpn. J. Appl. Phys.*, L307 (2007).
95. S. Nakamura, Current status of GaN-based solid-state lighting, *MRS Bull.*, 34, 101 (2009).
96. W. Schockley and H. J. Queisser, Detailed balance limit of efficiency of p-n junction solar cells, *J. Appl. Phys.*, 32, 510 (1961).
97. C. H. Henry, Limiting efficiencies of ideal single and multiple energy gap terrestrial solar cells, *J. Appl. Phys.*, 51, 4494 (1980).
98. S. M. Vernon, V. E. Haven, S. P. Tobin, and R. G. Wolfson, Metalorganic chemical vapor deposition of GaAs on Si for solar cell applications, *J. Cryst. Growth*, 77, 530 (1986).
99. J. F. Geisz, D. J. Friedman, J. S. Ward, A. Duda, W. J. Olavarria, T. E. Moriarty, J. T. Kiehl, M. J. Romero, A. G. Norman, and K. M. Jones, 40.8% efficient inverted triple-junction solar cell with two independently metamorphic junctions, *Appl. Phys. Lett.*, 93, 123505 (2008).
100. R. R. King, D. C. Law, K. M. Edmondson, C. M. Fetzer, G. S. Kinsey, H. Yoon, R. A. Sherif, and N. H. Karam, 40% efficient metamorphic GaInP/GaInAs/Ge multijunction solar cells, *Appl. Phys. Lett.*, 90, 183516 (2007).
101. M. Stan, D. Aiken, B. Cho, A. Cornfeld, V. Ley, P. Patel, P. Sharps, and T. Varghese, High-efficiency quadruple junction solar cells using OMVPE with inverted metamorphic device structures, *J. Cryst. Growth*, 312, 1370 (2010).
102. D. Aiken, I. Aeby, B. Clevenger, F. Newman, P. Patel, P. Sharps, M. Stan, T. Varghese, and J. Wood, Concentrator solar cell production capability and laboratory results at Emcore, in *6th International Conference on Concentrating Photovoltaic Systems* (2010), p. 7.
103. J. Simon, S. Tomasulo, P. J. Simmonds, M. Romeo, and M. L. Lee, Metamorphic InGaP on GaAs and GaP for wide-bandgap photovoltaic junctions, in *2010 35th IEEE Photovoltaic Specialists Conference* (2010), p. 002106.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Appendix I: Bandgap Engineering Diagrams

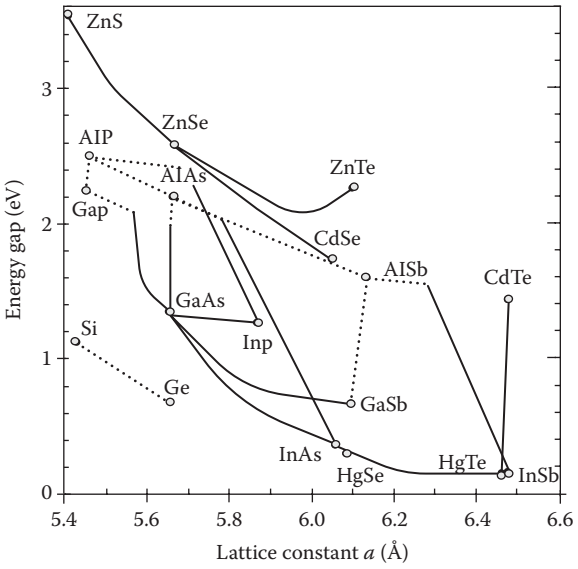


FIGURE I.1
Energy gap as a function of lattice constant *a* for cubic semiconductors. Room temperature values are given. Dashed lines indicate an indirect gap.

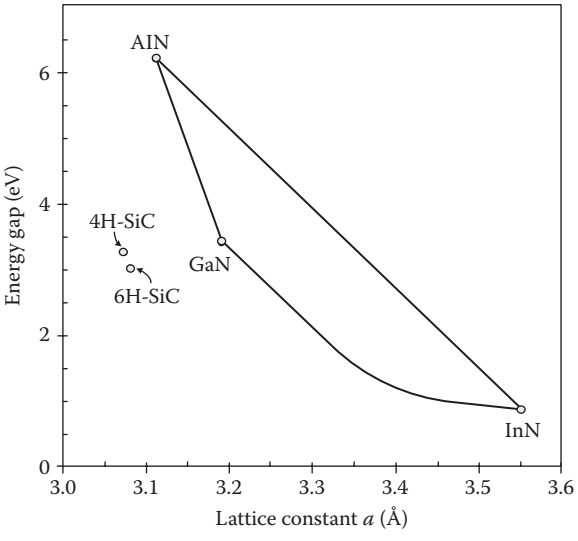


FIGURE I.2
Energy gap as a function of lattice constant *a* for hexagonal semiconductors. Room temperature values are given. Sapphire, a commonly used substrate material for III-nitrides, has room temperature lattice constants of $a = 0.47592$ nm and $c = 1.29916$ nm.[1]

Reference

1. Y. V. Shvyd'ko, M. Lucht, E. Gerdau, M. Lerche, E. E. Alp, W. Sturhahn, J. Sutter, and T. S. Toellner, Measuring wavelengths and lattice constants with the Mössbauer wavelength standard, *J. Synchrotron Rad.*, 9, 17 (2002).

Appendix II: Electronic Properties of Semiconductors

The applications of heteroepitaxial materials are ultimately driven by their electronic properties. Some of the most important such properties are tabulated below (Tables II.1 and II.2), including the minimum energy gaps at 300 K and 0 K with their character (direct or indirect), the carrier mobilities and effective masses at 300 K, and the relative permittivity.

TABLE II.1

Electronic Properties of Cubic Semiconductors

	Energy Gap (eV)		Indirect (I)/ Direct (D)	Mobility at 300 K		Effective Masses		Permittivity
	300 K	0 K		μ_n	μ_p	m_n^*/m_0	m_p^*/m_0	$\varepsilon/\varepsilon_0$
C[1]	5.47	5.48	I	1800	1200	0.2	0.25	5.7
Si[1]	1.12	1.17	I	1450	500	$0.98^{l}/0.19^{t}$	$0.16^{lh}/0.49^{hh}$	11.9
Ge[1]	0.66	0.74	I	3900	1900	$1.64^{l}/0.082^{t}$	$0.041^{lh}/0.28^{hh}$	16.0
AlAs[1]	2.36	2.23	I	180		0.11	0.22	10.1
AlP[1]	2.42	2.51	I	60	450	0.212	0.145	9.8
AlSb[1]	1.58	1.68	I	200	450	0.12	0.98	14.4
BN[1]	6.4	—	I	200	500	0.26	0.36	7.1
BP[1]	—	—	I	40	500	0.67	0.042	11
GaAs[1]	1.42	1.52	D	8,000	400	0.063	$0.076^{lh}/0.5^{hh}$	12.9
GaP[1]	2.26	2.34	I	110	75	0.82	0.60	11.1
GaSb[1]	0.72	0.81	D	5,000	850	0.042	0.40	15.7
InAs[1]	0.36	0.42	D	33,000	460	0.023	0.40	15.1
InP[1]	1.35	1.42	D	4,600	150	0.077	0.64	12.6
InSb[1]	0.17	0.23	D	80,000	1,250	0.0145	0.40	16.8
CdS[1]	2.5	—	D	—	—	0.14	0.51	5.4
CdSe[1]	1.70	1.85	D	800	—	0.13	0.45	10.0
CdTe[1]	1.56	—	D	1,050	100	—	—	10.2
HgSe[2]	–0.30	—	D	20,000		—	—	—
HgTe[2]	–0.15	—	D	25,000	—	—	—	—
ZnS[1]	3.66	3.84	D	600	—	0.39	0.23	8.4
ZnSe[2]	2.58	—	D	—	—	—	—	8.35
ZnTe[2]	2.26	—	D	340	100	—	—	12.67

Note: m_n^*/m_0 is the normalized conductivity effective mass for electrons and m_p^*/m_0 is the normalized conductivity effective mass for holes. t = transverse; l = longitudinal; lh = light hole; hh = heavy hole. The (–) sign in front of the value for the energy gap implies that the compound acts as a semimetal and therefore there exists an overlapping of the valence and conduction band.

TABLE II.2

Electronic Properties of Hexagonal Semiconductors

	Energy Gap (eV)		Indirect (I)/ Direct (D)	Mobility at 300 K		Effective Masses		Permittivity
	300 K	0 K		μ_n	μ_p	m_n^*/m_0	m_p^*/m_0	$\varepsilon/\varepsilon_0$
2H-SiC	3.33[3]	—	I	—	—	—	—	9.66
4H-SiC	3.2633	—	I	450[4]	1154	—	—	9.66
6H-SiC	3.023[3]	—	I	500[4]	101[4]	_	—	9.66[5]
AlN	6.03[6]	6.13[6]	D	—	—	0.27[7]	—	8.58
BN	6.1[9]	—	I	—	—	—	—	—
GaN[1]	3.44	3.50	D	400	10	0.27	0.8	10.4
InN	0.641[10]	0.690[11]	D	—	—	0.11[12]	—	—
CdS[1]	2.49	2.58[13]	D	350	40	0.20	0.7	9.1
ZnS[1]	3.66	3.84	D	600	—	0.39	0.23	8.4

Note: m_n^*/m_0 is the normalized conductivity effective mass for electrons and m_p^*/m_0 is the normalized conductivity effective mass for holes. t = transverse; l = longitudinal; lh = light hole; hh = heavy hole.

References

1. S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. (Hoboken, NJ: John Wiley & Sons, 2007).
2. L. I. Berger and B. R. Pamplin, Properties of semiconductors, in *CRC Handbook of Physics and Chemistry*, ed. D. R. Lide, 84th ed. (Boca Raton, FL: CRC Press, 2004), pp. 12–107.
3. S. Yoshida, Band structure of SiC, in *Properties of Silicon Carbide*, ed. G. L. Harris (Exeter, England: Short Run Press Ltd., 1995).
4. G. L. Harris, H. S. Henry, and A. Jackson, Carrier mobilities and concentrations in SiC, in *Properties of Silicon Carbide*, ed. G. L. Harris (Exeter, England: Short Run Press Ltd., 1995), pp. 63–68.
5. L. Patrick and W. J. Choyke, Static dielectric constant of SiC, *Phys. Rev. B*, 2, 2255 (1970).
6. Q. Guo and A. Yoshida, Temperature dependence of band gap change in InN and AlN, *Jpn. J. Appl. Phys.*, 33, 2453 (1994).
7. M. Suzuki and T. Uenoyama, First-principles calculations of effective-mass parameters of AlN and GaN, *Phys Rev. B*, 52, 8132 (1995).
8. O. Ueda and S. J. Pearton, *Materials and Reliability Handbook for Semiconductor Optical and Electron Devices* (New York: Springer Science + Business Media, 2013).
9. S. L. Rumyantsev, M. E. Levinshtein, A. D. Jackson, S. N. Mohammmad, G. L. Harris, M. G. Spencer, and M. S. Shur, Boron nitride, in *Properties of Advanced Semiconductor Materials GaN, AlN, InN, BN, SiC, SiGe*, ed. M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur (New York: John Wiley & Sons, 2001), p. 67.
10. J. Wu, W. Walukiewicz, W. Shan, K. M. Yu, J. W. Ager III, S. X. Li, E. E. Haller, H. Lu, and W. Schaff, Temperature dependence of the fundamental band gap of InN, *J. Appl. Phys.*, 94, 4457 (2003).
11. W. Walukiewicz, J. W. Ager III, K. M. Yu, Z. Liliental-Weber, J. Wu, S. X. Li, R. E. Jones, and J. D. Denlinger, Structure and electronic properties of InN and In-rich group III-nitride alloys, *J. Phys. D Appl. Phys.*, 39, R83 (2006).
12. H. Morkoc, *Handbook of Nitride Semiconductors and Devices*, Vol. 1: *Materials Properties, Physics and Growth* (Weinheim: Wiley-VCH Verlag, 2008).
13. M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 2nd ed. (Berlin: Springer-Verlag, 1998 and 1999).

Appendix III: Lattice Constants and Coefficients of Thermal Expansion

Tables III.1 and III.2 give the lattice constants and thermal expansion coefficients of selected cubic and hexagonal crystals. The linear thermal coefficient of expansion (TCE) α is defined as

$$\alpha \equiv \frac{1}{a}\frac{\partial a}{\partial T}. \tag{III.1}$$

The variation of the lattice constant with temperature can also be fit to a polynomial:

$$\frac{\Delta a}{a} = A + BT + CT^2 + DT^3, \tag{III.2}$$

TABLE III.1

Lattice Constants and Thermal Expansion Coefficients for Cubic Semiconductor Crystals

	a (300 K) (Å)	α (300 K) (10^{-6} K^{-1})	α (600 K) (10^{-6} K^{-1})	α (1000 K) (10^{-6} K^{-1})
C	3.56684[1]	1.0[2]	2.8[2]	4.4[2]
Si	5.43108[3]	2.6[2]	3.7[2]	4.4[2]
Ge	5.6576[4]	5.7[2]	6.7[2]	7.6[2]
α-Sn	6.4894[5]	4.7[5]	—	—
SiC (3C)	4.3596[6,7]	—	—	—
BN	3.615[8]	1.8[9]	3.7[9]	5.9[9]
BP	4.538[10]	—	—	—
BAs	4.777[10]	—	—	—
AlP	5.467[11]	—	—	—
AlAs	5.660[12]	—	—	—
AlSb	6.1357[13]	4.4[2]	—	—
GaP	5.4512[14]	4.7[9]	5.8[9]	—
GaAs	5.65325[15]	5.7[9]	6.7[9]	—
GaSb	6.0960[16]	6.1[2]	7.3[2]	—
InP	5.8690[13]	4.75[17]	—	—
InAs	6.0584[18]	5.19[19]	—	—
InSb	6.4794[20]	5.0[2]	6.1[2]	—
BeS	4.865[10]	—	—	—
BeSe	5.139[10]	—	—	—
BeTe	5.626[10]	—	—	—
ZnS	5.4105[21]	7.1[9]	8.6[9]	10.5[9]
ZnSe	5.6687[22]	7.1[9]	10.1[9]	—
ZnTe	6.1041[23]	8.8[9]	10.0[9]	—
CdTe	6.481[24]	5.0[2]	5.4[2]	—
β-HgS	5.851[21]	—	—	—
HgSe	6.084[21]	—	—	—
HgTe	6.461[25]	5.1[9]	—	—

TABLE III.2

Lattice Constants and Thermal Expansion Coefficients for Hexagonal Crystals

	a (Å)	c (Å)	α_a (300 K) (10^{-6} K^{-1})	α_c (300 K) (10^{-6} K^{-1})	α_a (600 K) (10^{-6} K^{-1})	α_c (600 K) (10^{-6} K^{-1})
α-Al_2O_3	4.7592[26]	12.9916[26]	4.3	3.9	5.6	7.4
SiC (2H)	3.0763[6]	5.0480[6]	—	3.0[27]	—	3.4[27]
SiC (4H)	3.0730[23]	10.05[3]	3.3[28]	3.2[28]	4.2[28]	3.8[28]
SiC (6H)	3.0806[6]	15.1173[6]	3.4[27]	3.3[27]	4.2[27]	3.9[27]
AlN	3.112[29]	4.9788[30]	4.4[30]	3.5[30]	—	—
BN	2.534[31]	4.191[31]				
GaN	3.1882(4)[32,33]	5.1855(5)[32,33]	3.1[34]	2.8[34]	4.7[34]	4.2[34]
InN	3.533[35]	5.693[35]	3.4[36]	2.7[36]	5.7[36]	3.7[36]
ZnS	3.8140	6.2576	8.4[37]	4.1[37]	9.9[37]	5.5[37]
ZnTe	4.27	6.99	—	—	—	—
CdS	4.1363[30]	6.7154[30]	4.3[30]	2.8[30]	—	—
CdSe	4.3007[30]	7.0113[30]	4.1[30]	2.8[30]	—	—
CdTe	4.57	7.47	—	—	—	—

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 1.562(3) is 1.562 ± 0.003.

TABLE III.3

Temperature Dependence of Thermal Expansion Coefficients for Cubic Crystals

	A (10^{-2})	B (10^{-6} K^{-1})	C (10^{-9} K^{-2})	D (10^{-12} K^{-3})	Temp. Range
C[2]	−0.010	−0.591	3.32	−0.5544	(25–1650K)
Si[2]	−0.071	1.887	1.934	−0.4544	(293–1600K)
Ge[2]	−0.1533	4.636	2.169	−0.4562	(293–1200K)
α-Sn[2]	−0.525	13.54	15.87	−2.896	(100–500K)
BN[9]	−0.0013	−1.278	4.911	−0.8635	(293–1300K)
AlSb[2]	−0.049	−2.997	22.43	−22.34	(40–350K)
GaP[9]	−0.110	2.611	4.445	−2.023	(293–850K)
GaAs[9]	−0.147	4.239	2.916	−0.936	(200–1000K)
GaAs[38–42]	−0.155	4.411	2.70	−0.667	(293–643K)
GaSb[2]	−0.138	3.051	66.02	−3.380	(100–800K)
InSb[2]	−0.099	1.249	8.773	−5.260	(50–750K)
ZnS[9]	−0.0863	−3.386	30.18	−29.21	(60–335K)
ZnS[43–45]	−0.2173	6.628	2.15	−0.333	(293–820K)
ZnSe[37,46–48]	−0.1566	3.044	8.25	−3.33	(293–870K)
ZnTe[9]	−0.200	5.104	6.811	−3.104	(100–725K)
CdTe[2]	−0.0980	1.624	7.176	−4.445	(100–700K)
HgTe[9]	−0.504	9.772	42.66	−59.22	(50–300K)

Note: $\Delta a/a = A + BT + CT^2 + DT^3$, where T is the absolute temperature.[a]

[a] These polynomial coefficients are absolute, so that $\Delta a/a$ is absolute. In the first edition, the coefficients were given such that $\Delta a/a$ was in percent.

where $\Delta a/a$ is measured with respect to 300K, and T is the absolute temperature in Kelvin. Thus, at a temperature T, the relaxed lattice constant for the cubic crystal is given by

$$a(T) = a(300K)\left[1 + A + BT + CT^2 + DT^3\right]. \tag{III.3}$$

From the polynomial characteristic, the linear coefficient of thermal expansion may be determined as a function of temperature from

$$\alpha = \frac{1}{a}\frac{\partial a}{\partial T} = B + 2CT + 3DT^2. \tag{III.4}$$

The constants A, B, C, and D for cubic crystals are provided in Table III.3. In the cases of GaAs and ZnS, two polynomials are given for two different temperature ranges.

References

1. W. Kaiser and W. L. Bond, Nitrogen, a major impurity in common type I diamond, *Phys. Rev.*, 115, 857 (1959).
2. Y. S. Touloukian, R. K. Kirby, R. E. Taylor, and P. D. Desai, eds., *Thermophysical Properties of Matter*, Vol. 12: *Thermal Expansion, Metallic Elements and Alloys* (New York: Plenum, 1975).
3. E. R. Cohen and B. N. Taylor, The 1986 adjustment of the fundamental physical constants, report of the Committee on Data for Science and Technology of the International Council of Scientific Unions (CODATA) Task Group on Fundamental Constants, CODATA Bulletin 63 (Elmsford, NY: Pergamon, 1986).
4. J. Donahue, *The Structure of the Elements* (New York: J. Wiley & Sons, 1974).
5. J. Thewlis and A. R. Davey, Thermal expansion of grey tin, *Nature*, 174, 1011 (1959).
6. A. Taylor and R. M. Jones, Silicon Carbide—a high temperature semiconductor, in J. R. O'Connor and J. Smiltens, (eds), *Proceedings of Conference on Silicon Carbide*, Boston, MA (Oxford: Pergamon Press, 1960), p. 147.
7. Y. M. Tairov and V. F. Tsvetkov, Progress in controlling the growth of polytypic crystals, in P. Krishna, (ed.), *Crystal Growth and Characterization of Polytype Structures*, (Oxford: Pergamon Press, 1983), p. 111.
8. R. C. Weast, ed., *Handbook of Chemistry and Physics*, 56th ed. (Cleveland: CRC Press, 1975).
9. Y. S. Touloukian, R. K. Kirby, R. E. Taylor, and T. Y. R. Lee, eds., *Thermophysical Properties of Matter*, Vol. 13: Thermal Expansion, Nonmetallic Solids (New York: Plenum, 1977).
10. R. C. Weast, ed., *CRC Handbook of Chemistry and Physics*, 58th ed. (West Palm Beach, FL: CRC Press, 1977).
11. C. C. Wang, M. Zaheeruddin, and L. H. Spinar, Preparation and properties of aluminum phosphide, *J. Inorg. Nucl. Chem.*, 25, 326 (1961).
12. W. B. Pearson, *A Handbook of Lattice Spacings and Structure of Metals and Alloys* (Oxford: Pergamon Press, 1967).
13. G. Giesecke and H. Pfister, Prazionbestimmung der Gitterkonstanten von $A_{III}B_V$—Verbingungen, *Acta Cryst.*, 11, 369 (1958).
14. H. C. Casey Jr. and F. A. Trumbore, Single crystal electroluminescent materials, *Mater. Sci. Eng.*, 6, 69 (1970).
15. E. M. H. Driscoll, A. F. W. Willoughby, J. B. Mullin, and B. W. Straughan, in *Gallium Arsenide and Related Compounds*, Vol. 24 in Institute of Physics Conference Series (London: Institute of Physics, 1975), p. 275.

16. M. E. Straumanis and C. D. Kim, Lattice parameters, thermal expansion coefficients, phase width, and perfection of the structure of GaSb and InSb, *J. Appl. Phys.*, 36, 3822 (1965).
17. I. Kudman and R. J. Paff, Thermal expansion of $In_xGa_{1-x}As$ alloys, *J. Appl. Phys.*, 43, 3760 (1972).
18. J. V. Ozolin'sh, G. K. Averkieva, A. F. Ilvin'sh, and N. A. Goryunova, An x-ray study of some A^3B^5 compounds which display deviations from stoichiometry, *Sov. Phys. Cryst.* (Engl. transl.), 7, 691 (1963).
19. M. Neuberger, *Handbook of Electronic Materials*, Vol. 2: *III-V Semiconductor Compounds* (New York: Plenum, 1971).
20. J. C. Jamieson and H. H. Demarest, A note on the compression of cubic ZnS, *J. Phys. Chem. Solids*, 41, 963 (1980).
21. W. L. Roth, in *Physics and Chemistry of II-VI Compounds*, ed. M. Aven and J. S. Prener, p. 134. (Amsterdam: North-Holland Publishing Co., 1975).
22. B. Greenberg, Private communication.
23. R. W. G. Wyckoff, *Crystal Structure* (New York: John Wiley & Sons, 1963), p. 108.
24. J. C. Woolley and B. Ray, Solid solution in $A^{II}B^{VI}$ tellurides, *J. Phys. Chem. Solids*, 13, 151 (1960).
25. R. T. Delves and B. Lewis, Zinc blende type HgTe-MnTe solid solutions. I, *J. Phys. Chem. Solids*, 24, 549 (1963).
26. Y. V. Shvyd'ko, M. Lucht, E. Gerdau, M. Lerche, E. E. Alp, W. Sturhahn, J. Sutter, and T. S. Toellner, Measuring wavelengths and lattice constants with the Mössbauer wavelength standard, *J. Synchrotron Rad.*, 9, 17 (2002).
27. Z. Li and R.C. Bradt, Thermal expansion and thermal expansion anisotropy of *SiC polytypes*, *J. Am. Cer. Soc.*, 70, 445 (1987).
28. Z. Li and R.C. Bradt, Thermal expansion of the hexagonal (4H) polytype of *SiC*, *J. Appl. Phys.*, 60, 612 (1986).
29. C. Kim, I. K. Robinson, J. Myoung, K. Shim, M.-C. Yoo, and K. Kim, Critical thickness of GaN thin films on sapphire (0001), *Appl. Phys. Lett.*, 69, 2358 (1996).
30. H. Iwanaga, A. Kunishige, and S. Takeuchi, Anisotropic thermal expansion in wurtzite-type crystals, *J. Mater. Sci.*, 35, 2451 (2000).
31. K. Shimada, T. Sota, and K. Suzuki, First-principles study on electronic and elastic properties of BN, AlN, and GaN, *J. Appl. Phys.*, 84, 4951 (1998).
32. S. Poroski, Bulk and homoepitaxial GaN-growth and characterisation, *J. Cryst. Growth*, 189–190, 153 (1998).
33. M. Leszczynski, H. Teisseyre, T. Suski, I. Grzegory, M. Bockowski, J. Jun, K. Pakula, J. M. Baranowski, C. T. Foxon, and T. S. Cheng, Lattice parameters of gallium nitride, *Appl. Phys. Lett.*, 69, 73 (1996).
34. M. Leszczynski, T. Suski, H. Teisseyre, P. Perlin, I. Grzegory, J. Jun, S. Porowski, and T. D. Moustakas, Thermal expansion of gallium nitride, *J. Appl. Phys.*, 76, 4909 (1994).
35. A. U. Shelag and V. A. Savastenko, Determination of elastic constants of hexagonal crystals from measured values of dynamic atomic displacements, *Izv. Akad. Nauk. BSSR Ser. Fiz. Mat. Nauk.*, 3, 126 (1976).
36. K. Wang and R. R. Reeber, Thermal expansion and elastic properties of InN, *Appl. Phys. Lett.*, 79, 1602 (2001).
37. S.Q. Wang, First-principles study of the anisotropic thermal expansion of wurtzite ZnS, *Appl. Phys. Lett.*, 88, 061902 (2006).
38. V. Grossman, M. Behringer, H. Heinke, and D. Hommel, Temperature-dependent measurements on ZnSe heterostructures by high-resolution x-ray diffraction, *J. Cryst. Growth*, 184–185, 100 (1998).
39. M. Leszczynski and J. F. Walker, Thermal expansion of gallium arsenide layers grown by molecular beam epitaxy at low temperatures, *Appl. Phys. Lett.*, 62, 1484 (1993).
40. R. Feder and I. Light, Precision thermal expansion measurements of semi-insulating GaAs, *J. Appl. Phys.*, 39, 4870 (1968).
41. E. D. Pierron, D. L. Parker, and J. B. McNeely, Coefficient of expansion of GaAs, GaP, and Ga(As,P) compounds from –62°C to 200°C, *J. Appl. Phys.*, 38, 4669 (1967).

42. S. I. Novikova, Investigation of thermal expansion of GaAs and ZnSe, *Sov. Phys. Sol. State*, 3, 129 (1961).
43. C. Giannini, L. Tapfer, T. Peluso, N. Lovergine, and L. Vasanelli, Residual and thermal strain of ZnS epitaxial layers grown on (100)-GaAs by vapour phase epitaxy, *J. Phys. D*, 28, A125 (1995).
44. J. S. Browder and S. S. Ballard, Low temperature thermal expansion measurements on optical materials, *Appl. Optics*, 8, 793 (1969).
45. R. B. Roberts, G. K. White, and T. M. Sabine, Thermal expansion of zinc sulfide: 300–1300K, *Aust. J. Phys.*, 34, 701 (1981).
46. S. S. Ballard, S. E. Brown, and J. S. Browder, Measurements of the thermal expansion of six optical materials, from room temperature to 250°C, *Appl. Optics*, 17, 1152 (1978).
47. H. P. Singh and B. Dayal, X-ray determination of the thermal expansion of zinc selenide, *Phys. Status Solidi B*, 23, K93 (1967).
48. L. S. Ladd, R. Clark, and D. Roy, U.S. Air Force Technical Report AL-TDR-64-9 (1964).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Appendix IV: Elastic Constants

Tables IV.1 through IV.5 provide the elastic stiffness constants of cubic and hexagonal crystals. Tables IV.6 and IV.7 give values of the elastic moduli.

TABLE IV.1

Elastic Stiffness Constants of Cubic Semiconductor Crystals at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

	C_{11}	C_{12}	C_{44}
C[1]	107.6	12.52(23)	57.74(14)
Si[2]	160.1	57.8	80.0
Ge[3]	124.0	41.3	68.3
α-Sn[4]	69.0	29.3	36.2
SiC (3C)[5]	352	120	232.9
AlN (ZB)[6]	322	156	138
AlP[7]	132	63.0	61.5
AlAs[7]	125	53.4	54.2
AlSb[8,9]	87.69(20)	43.41(20)	40.76(8)
GaN (ZB)[6]	325	142	147
GaP[10]	140.50(28)	62.03(24)	70.33(7)
GaAs[11]	118.4(3)	53.7(16)	59.1(2)
GaSb[12]	88.50	40.40	43.30
InP[13]	102.2	57.6	46.0
InAs[14]	83.29	45.26	39.59
InSb[15]	65.92(5)	35.63(6)	29.96(3)
ZnS[16]	104.62(5)	65.33(6)	46.50(12)
ZnSe[17]	87.2(8)	52.4(8)	39.2(4)
ZnTe[18]	71.3	40.7	31.2
CdTe[19]	53.3	36.5	20.44
β-HgS[20]	81.3	62.2	26.4
HgSe[20]	69.0	51.9	23.3
HgTe[21]	53.61	36.60	21.23

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 12.52(23) represents 12.52 ± 0.023.

TABLE IV.2

Elastic Stiffness Constants of 4H- and 6H-SiC at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

Elastic Constants	4H-SiC (Kamitani et al.[22])	6H-SiC (Kamitani et al.[22])
C_{11}	507(4)	501(4)
C_{12}	111(5)	111(5)
C_{13}	—	52(9)
C_{33}	547(7)	553(4)
C_{44}	159(4)	163(4)
C_{66}	198	195

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 507(4) represents 507 ± 4.

TABLE IV.3

Elastic Stiffness Constants of Wurtzite GaN at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

Elastic Constants	Recommended Values	Polian et al.[23]	Deger et al.[24]	Deguchi et al.[25]	Davydov et al.[26]	Savastenko and Sheleg[27]
C_{11}	353	390(15)	370	373	315	296
C_{12}	135	145(20)	145	141	118	120
C_{13}	104	106(20)	110	80.4	96	158
C_{33}	367	398(20)	390	387	324	267
C_{44}	91	105(10)	90	93.6	88	24
C_{66}	110	123(10)	112	118	99	88

Note: The number in parentheses represents the uncertainty in the last digit; i.e., 390(15) represents 390 ± 15.

TABLE IV.4

Elastic Stiffness Constants of Wurtzite AlN at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

Elastic Constants	Recommended Values	Deger et al.[24]	Davydov et al.[26]	McNeil et al.[28]	Tsubouchi et al.[29]	Davydov and Solomonov[6]
C_{11}	397	410	419	411	345	369
C_{12}	145	140	177	149	125	145
C_{13}	113	100	140	99	120	120
C_{33}	392	390	392	389	395	395
C_{44}	118	120	110	125	125	96
C_{66}	128	135	121	131	131	112

TABLE IV.5

Elastic Stiffness Constants of Wurtzite InN at Room Temperature, in Units of GPa (1 GPa = 10^{10} dyn/cm^2)

Elastic Constants	Recommended Values	Sheleg and Savastenko[30]	Kim et al.[31]	Wright[32]	Marmalyuk et al.[33]	Chisholm et al.[34]
C_{11}	250	190	271	223	257	297.5
C_{12}	109	104	124	115	92	107.4
C_{13}	98	121	94	92	70	108.7
C_{33}	225	182	200	224	278	250.5
C_{44}	54	9.9	46	48	68	89.4
C_{66}	70	43	74	54	82	95

TABLE IV.6

Elastic Moduli of Cubic Semiconductor Crystals at 300 K

	G(GPa)	E(001) (GPa)	ν(001)	Y(001) (GPa)	R^B(001)
C	47	105	0.104	117	0.23
Si	51	129	0.265	176	0.72
Ge	41	103	0.25	138	0.67
α-Sn	19.8	52	0.30	73	0.85
SiC (3C)	116	290	0.25	390	0.68
AlN (ZB)	83	220	0.33	330	0.97
AlP	34	91	0.32	135	0.95
AlAs	36	93	0.30	133	0.85
AlSb	22	59	0.33	88	0.99
GaN (ZB)	92	240	0.30	340	0.87
GaP	39	102	0.31	148	0.88
GaAs	32	85	0.31	124	0.91
GaSb	24	63	0.31	92	0.91
InP	22	61	0.36	95	1.13
InAs	19.0	51	0.35	79	1.09
InSb	15.1	41	0.35	63	1.08
ZnS	19.6	54	0.38	88	1.25
ZnSe	17.4	48	0.38	77	1.20
ZnTe	15.3	42	0.36	66	1.14
CdTe	8.4	24	0.41	40	1.37
β-HgS	9.6	27	0.43	48	1.53
HgSe	8.6	24	0.43	43	1.50
HgTe	8.5	24	0.41	40	1.37

Note: These values were calculated using the elastic stiffness constants of Table IV.1; see that table for references.

TABLE IV.7

Elastic Moduli of Hexagonal Semiconductor Crystals at 300 K

	$E(001)$ (GPa)	$\nu(0001)$	$Y(001)$ (GPa)	$R^B(0001)$
6H-SiC	540	0.085	602	0.19
GaN[23]	320	0.21	430	0.57
AlN	340	0.21	480	0.58
InN	171	0.27	270	0.87

References

1. M. H. Grimsditch and A. K. Ramdas, Brillouin scattering in diamond, *Phys. Rev. B*, 11, 3139 (1975).
2. S. P. Nikanorov, Yu. A. Burenkov, and A. V. Stepanov, Elastic properties of Si, *Sov. Phys. Solid State* (Engl. transl.), 13, 2516 (1972).
3. Yu. A. Burenkov, S. P. Nikanorov, and A. V. Stepanov, Elastic properties of germanium, *Sov. Phys. Solid State* (Engl. transl.), 12, 1940 (1971).
4. D. L. Price, J. M. Rowe, and R. M. Nicklow, Lattice dynamics of grey tin and indium antimonide, *Phys. Rev. B*, 3, 1268 (1971).
5. K. B. Tolpygo, Optical, elastic and piezoelectric properties of ionic and valence crystals with the ZnS type lattice, *Sov. Phys. Solid State*, 2, 2367 (1961).
6. S. Yu. Davydov and A. V. Solomonov, Elastic properties of gallium and aluminum nitrides, *Tech. Phys. Lett.*, 25, 601 (1999) (transl. from *Pis'ma Zh. Tekh. Fiz.*, 25, 23 [1999]).
7. J. D. Wiley, *Semiconductors and Semimetals*, Vol. 10, ed. R. K. Willardson and A. C. Beer (New York: Academic Press, 1975).
8. D. I. Bolef and M. Menes, Elastic constants of single-crystal aluminum antimonide, *J. Appl. Phys.*, 31, 1426 (1960).
9. R. Weil, Correction to the elastic constants of AlSb, *J. Appl. Phys.*, 43, 4271 (1972).
10. Y. K. Yogurtcu, A. J. Miller, and G. A. Saunders, Pressure dependence of elastic behaviour and force constants of GaP, *J. Phys. Chem. Solids*, 42, 49 (1981).
11. R. I. Cottam and G. A. Saunders, The elastic constants of GaAs from 2K to 320K, *J. Phys., C6*, 2105 (1973).
12. G. Simmons and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook* (Cambridge, MA: MIT Press, 1971).
13. F. S. Hinkernell and W. R. Gayton, Elastic constants of single-crystal indium phosphide, *J. Appl. Phys.*, 37, 462 (1966).
14. D. Gerlich, Elastic constants of single-crystal indium phosphide, *J. Appl. Phys.*, 34, 2915 (1963).
15. I. O. Baskin and G. I. Pereseda, Influence of hydrostatic pressures up to 17 kbar on the elastic moduli of InSb, *Sov. Phys. Solid State* (Engl. transl.), 16, 2058 (1975).
16. I. M. Sil'vestrova, I. B. Kobyakov, and A. A. Shternberg, Elastic and piezoelectric properties of single crystals of sphalerite (ZnS) grown by the hydrothermal method, *Sov. Phys. Techn. Phys.* (Engl. transl.), 25, 1433 (1980).
17. C. G. Hodgins and J. C. Irwin, Elastic constants of ZnSe, *Phys. Status Solidi, A28*, 647 (1975).
18. D. Berlincourt, H. Jafe, and L. R. Shiozawa, Electroelastic properties of the sulfides, selenides, and tellurides of zinc, *Phys. Rev.*, 129, 1009 (1963).
19. Yu. Yekilov and A. P. Rusakov, Elastic constants and lattice dynamics of some AB compounds, *Sov. Phys. Solid State* (Engl. transl.), 13, 956 (1971).

20. K. Kumazaki, Elastic properties and iconicity of zero-gap semiconductors, *Phys. Status Solidi A*, 33, 615 (1976).
21. R. I. Cottam and G. A. Saunders, The elastic behavior of mercury telluride, *J. Phys. Chem. Solids*, 36, 1987 (1975).
22. K. Kamitani, M. Grimsditch, J. C. Nipko, C.-K. Loong, M. Okada, and I. Kimura, The elastic constants of silicon carbide: A Brillouin-scattering study of 4H and 6H SiC single crystals, *J. Appl. Phys.*, 82, 3152 (1997).
23. A. Polian, M. Grimsditch, and I. Grzegory, Elastic constants of gallium nitride, *J. Appl. Phys.*, 79, 3343 (1996).
24. B. Deger, E. Born, H. Angerer, O. Ambacher, M. Stutzmann, J. Hornsteiner, E. Riha, and G. Fischerauer, Sound velocity of $Al_xGa_{1-x}N$ thin films obtained by surface acoustic-wave measurements, *Appl. Phys. Lett.*, 72, 2400 (1998).
25. T. Deguchi, D. Ichiryu, K. Toshikawa, K. Segiguchi, T. Sota, R. Matsuo, T. Azuhata, M. Yamaguchi, T. Yagi, S. Chichibu, and S. Nakamura, Structural and vibrational properties of GaN, *J. Appl. Phys.*, 86, 1860 (1999).
26. V. Yu. Davydov, Yu. E. Kitaev, I. N. Goncharuk, A. N. Smirnov, J. Graul, O. Semchinova, D. Uffmann, M. B. Smirnov, A. P. Mirgorodsky, and R. A. Evarestov, Phonon dispersion and Raman scattering in hexagonal GaN and AlN, *Phys. Rev. B*, 58, 12899 (1998).
27. V. A. Savastenko and A. U. Sheleg, Study of the elastic properties of gallium nitride, *Phys. Status Solidi A*, 48, K135 (1978).
28. L. E. McNeil, M. Grimsditch, and R. H. French, Vibrational spectroscopy of aluminum nitride, *J. Am. Ceram. Soc.*, 76, 1132 (1993).
29. K. Tsubouchi, K. Sugai, and N. Mikoshiba, AlN material constants evaluation and SAW properties on AlN/Al_2O_3 and AlN/Si, in *1981 Ultrasonics Symposium Proceedings*, ed. B. R. McAvoy (New York: IEEE, 1981), p. 375.
30. A. U. Sheleg and V. A. Savastenko, Determination of elastic constants of hexagonal crystals from measured values of dynamic atomic displacements, *Izv. Akad. Nauk. BSSR Ser. Fiz. Mat. Nauk.*, 3, 126 (1976).
31. K. Kim, W. R. L. Lambrecht, and B. Segall, Elastic constants and related properties of tetrahedrally bonded BN, AlN, GaN, and InN, *Phys. Rev. B*, 53, 16310 (1996).
32. A. F. Wright, Elastic properties of zinc-blende and wurtzite AlN, GaN, and InN, *J. Appl. Phys.*, 82, 2833 (1997).
33. A. A. Marmalyuk, R. K. Akchurin, and V. A. Gorbylev, Evaluation of elastic constants of AlN, GaN, and InN, *Inorg. Mater.*, 34, 691 (1998) (transl. of *Neorganicheskie Materialy*, 34, 833 [1998]).
34. J. A. Chisholm, D. W. Lewis, and P. D. Bristowe, Classical simulations of the properties of group-III nitrides, *J. Phys. Cond. Mater.*, 11, L235 (1999).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Appendix V: The Critical Layer Thickness in Cubic Semiconductors

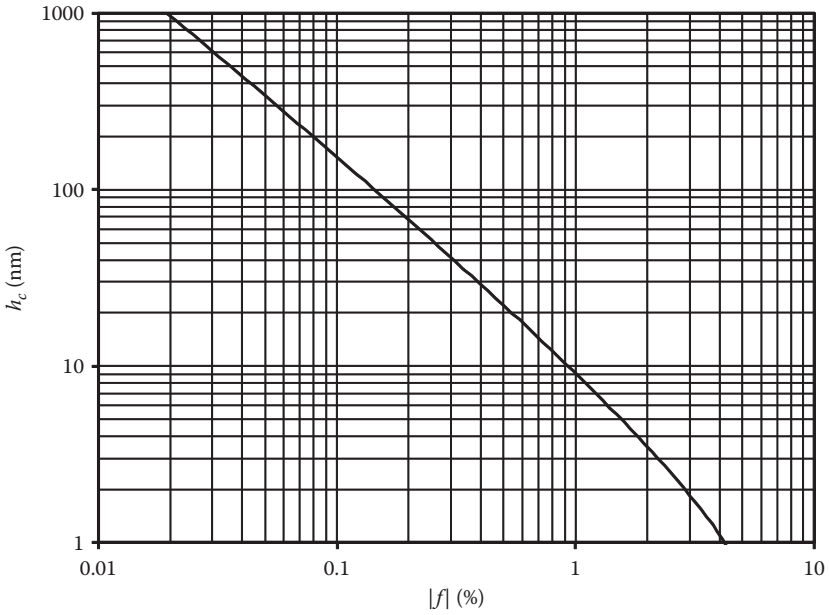


FIGURE V.1
Critical layer thickness as a function of lattice mismatch for cubic semiconductors calculated using the model of Matthews and Blakeslee [1] with the material parameters for GaAs.

Reference

1. J. W. Matthews and A. E. Blakeslee, Defects in epitaxial multilayers I. Misfit dislocations, *J. Cryst. Growth*, 27, 118 (1974).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Appendix VI: Crystallographic Etches

Table VI.1 gives selected crystallographic etches for some semiconductor materials.

TABLE VI.1

Crystallographic Etches for Semiconductors

Semiconductor	Etchant	Remarks
Si	1 mL HF 3 mL HNO_3 10 mL CH_3COOH	Dash etch,[1] 8 h
	1 mL HF 1 mL CrO_3 (5 M in H_2O)	Sirtl etch,[2] for (111)Si, 5 min
	2 mL HF 1 mL $K_2Cr_2O_7$ (0.15 M in H_2O)	Secco etch,[3] for (001)Si and (111), 5 min
	60 mL HF 30 mL HNO_3 60 mL CH_3COOH (glacial) 60 mL H_2O 30 mL solution of 1 g CrO_3 in 2 mL H_2O 2 g $(CuNO_3)_2{:}3H_2O$	Wright etch,[4] for (001)Si and (111), 5 min, long shelf life
	2 mL HF 1 mL HNO_3 2 mL $AgNO_3$ (0.65 M in H_2O)	Silver etch, reveals stacking faults
Ge	50 mL HF 100 mL HNO_3 110 mL CH_3COOH (glacial) 330 mg I_2	Reveals dislocations and stacking faults[5]
GaAs	1 mL Br_2 100 mL CH_3OH	Distinguishes between (111)Ga and (111)As planes
	1 mL HF 2 mL H_2O 8 mg $AgNO_3$ 1 g CrO_3	For (001)GaAs and (011)
	A: 40 mL HF 40 mL H_2O 0.3 g $AgNO_3$ B: 40 mL H_2O 40 g CrO_3	A-B etch,[6,7] separate parts store indefinitely, mix in a 1:1 ratio before use
	1 g $K_3Fe(CN)_6$ in 50 mL H_2O 12 mL NH_4OH in 36 mL H_2O	Mix in a 1:1 ratio before use
	KOH	Molten KOH,[8,9] 250°C
InP	HBr HF	Pyramidal pits[10] on (001)
	HBr CH_3COOH	Rectangular pits[11] on (001)
InSb	2 mL HNO_3 1 mL HF 1 mL CH_3COOH (glacial)	Round pits[12] on (001) and (111)

(*Continued*)

TABLE VI.1 (CONTINUED)

Crystallographic Etches for Semiconductors

Semiconductor	Etchant	Remarks
ZnSe	1 mL Br_2 25 mL CH_3OH	Fast etch,[13,14] 6 s
GaN	1 mL H_3PO_4 4 mL H_2SO_4	230°C, 10 min[15]
	KOH	Molten KOH, 250°C

References

1. W. C. Dash, Copper precipitation on dislocations in silicon, *J. Appl. Phys.*, 27, 1193 (1956).
2. E. Sirtl and A. Adler, Chromsäure-Flussäure als spezifisches System zur Ätzgruben entwicklung auf Silizium, *Z. Met.*, 52, 529 (1961).
3. F. Secco d'Aragona, Dislocation etch for (100) planes in silicon, *J. Electrochem. Soc.*, 119, 948 (1972).
4. M. W. Jenkins, A new preferential etch for defects in silicon crystals, *J. Electrochem. Soc.*, 124, 757 (1977).
5. Q. Li, Y.-B. Jiang, H. Xu, S. Hersee, and S. M. Han, Heteroepitaxy of high-quality Ge on Si by nanoscale Ge seeds grown through a thin layer of SiO_2, Appl. *Phys. Lett.*, 85, 1928 (2004).
6. M. S. Abrahams and C. J. Buicchi, Etching of dislocations on the low index planes of GaAs, *J. Appl. Phys.*, 36, 2855 (1965).
7. G. H. Olsen and M. Ettenberg, Universal stain/etchant for interfaces in III-V compounds, *J. Appl. Phys.*, 36, 2855 (1965).
8. J. G. Grabmaier and C. B. Watson, *Phys. Status Solidi*, 32, K13 (1967).
9. J. Angilello, R. M. Potemski, and G. R. Woolhouse, Etch pits and dislocations in {100} GaAs wafer, *J. Appl. Phys.*, 46, 2315 (1975).
10. K. Akita, T. Kusunoki, S. Komiya, and T. Kotani, Observation of etch pits in InP by new etchants, *J. Cryst. Growth*, 46, 783 (1979).
11. K. Akita, T. Kusunoki, S. Komiya, and T. Kotani, Observation of etch pits in InP by new etchants, *J. Cryst. Growth*, 46, 783 (1979).
12. H. C. Gatos and M. C. Lavine, Dislocation etch pits on the {111} and surfaces of InSb, *J. Appl. Phys.*, 31, 743 (1960).
13. A. Kamata and H. Mitsuhashi, Characterization of ZnSSe on GaAs by etching and X-ray diffraction, *J. Cryst. Growth*, 142, 31 (1994).
14. X. G. Zhang, A. Rodriguez, X. Wang, P. Li, F. C. Jain, and J. E. Ayers, Complete removal of threading dislocations from mismatched layers by patterned heteroepitaxial processing, *Appl. Phys. Lett.*, 77, 2524 (2000).
15. Y. Ono, Y. Iyechika, T. Takada, K. Inui, and T. Matsye, Reduction of etch pit density on GaN by InGaN-strained SQW, *J. Cryst. Growth*, 189, 133 (1998).

Appendix VII: Tables for X-Ray Diffraction of Cubic Materials

Tables VII.1 through VII.6 contain x-ray diffraction data for Si, GaAs, and InP crystals with (001) and (111) orientations. The values included are the most important for the design and interpretation of x-ray diffraction experiments, for a number of *hkl* reflections. They include the Bragg angle θ_B; the inclination of the (*hkl*) planes with respect to the surface ϕ; the theoretical rocking curve full width at half maximums for $(\theta_B + \phi)$ and $(\theta_B - \phi)$ incidence, W_+ and W_-, respectively; the extinction depth t_{ext}; the absorption depth t_{abs}; the penetration depth t_p; and the magnitude of the structure factor $|F_h|$.

TABLE VII.1

X-Ray Diffraction Data for Si(001) Crystals (a = 0.543108 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
113	28.060	25.239	0.8	12.3	15.7	3.2	2.7	46
004	34.564	0	3.5	3.5	34.3	19.8	12.6	60
224	44.014	35.264	1.2	7.5	26.2	9.2	6.8	54
115	47.474	15.793	1.5	2.6	68.3	23.1	17.3	36
044	53.352	45.000	1.1	7.2	28.6	8.8	6.8	48
135	57.044	32.312	1.3	3.1	71.6	20.6	16.0	33
026	63.769	18.435	2.6	3.6	69.9	28.9	20.5	44
335	68.443	40.316	1.7	3.4	81.6	22.0	17.3	30
353	68.443	62.774	0.9	6.6	33.3	6.1	5.2	30
444	79.307	54.736	4.6	7.9	50.1	18.4	13.5	40

TABLE VII.2

X-Ray Diffraction Data for Si(111) Crystals (a = 0.543108 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
111	14.221	0	6.8	6.8	15.2	8.6	5.5	59
224	44.014	19.471	2.0	4.3	41.3	19.8	13.4	54
115	47.475	38.942	0.8	5.2	38.4	9.0	7.3	36
333	47.475	0	2.0	2.0	73.5	25.8	19.1	36
044	53.352	35.264	1.5	5.0	42.0	16.6	11.9	48
135	57.044	28.561	1.4	2.9	76.3	22.5	17.4	33
026	63.769	43.089	1.8	5.0	48.4	18.0	13.1	44
335	68.443	14.420	2.2	2.6	109.5	31.2	24.2	30
444	79.307	0	6.0	6.0	90	34.3	24.9	40

TABLE VII.3

X-Ray Diffraction Data for GaAs(001) Crystals (a = 0.56534 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
002	15.814	0	0.5	0.5	200.4	4.0	3.9	6
113	26.866	25.239	1.5	40.4	4.9	0.8	0.7	127
004	33.026	0	8.7	8.7	13.7	8.0	5.0	163
224	41.875	35.264	2.4	20.5	9.5	3.0	2.3	144
115	45.072	15.793	3.5	6.3	27.4	9.1	6.9	98
044	50.423	45.000	2.0	20.8	9.7	2.5	2.0	130
135	53.715	32.312	2.7	7.5	28.1	7.8	6.1	89
006	54.838	0	0.3	0.3	530.8	11.9	11.7	6
026	59.513	18.435	5.4	8.0	28.0	11.5	8.1	118
335	63.313	40.316	3.1	7.7	31.4	8.1	6.5	80
117	76.667	11.422	7.7	8.5	52.7	13.9	11.0	75

TABLE VII.4

X-Ray Diffraction Data for GaAs(111) Crystals (a = 0.56534 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
111	13.650	0	16.4	16.4	6.2	3.4	2.2	155
224	41.875	19.471	4.6	10.7	16.5	7.8	5.3	144
115	45.072	38.942	1.6	14.5	13.7	2.8	2.3	98
333	45.072	0	4.8	4.8	29.7	10.3	7.7	98
044	50.423	35.264	3.3	12.5	16.7	6.0	4.4	130
135	53.715	28.561	3.0	6.9	30.2	8.7	6.7	89
335	63.313	14.42	4.3	5.5	43.7	12.4	9.7	80
444	70.733	0	8.4	8.4	35.9	13.8	10.0	108

TABLE VII.5

X-Ray Diffraction Data for InP(001) Crystals (a = 0.58690 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
002	15.218	0	10.1	10.1	10.4	1.4	1.2	118
113	25.804	25.239	0.9	70.5	2.8	0.1	0.1	143
004	31.668	0	8.2	8.2	14.5	2.8	2.3	168
224	40.015	35.264	2.0	22.8	8.7	0.8	0.7	151
115	42.999	15.793	3.9	7.2	24.9	3.2	2.8	121
044	47.941	45.000	1.4	26.8	7.6	0.5	0.5	139
135	50.939	32.312	2.6	8.1	24.4	2.6	2.3	104
006	51.952	0	3.2	3.2	53.3	4.2	3.9	70
026	56.108	18.435	4.9	7.7	27.5	4.0	3.5	130
335	59.390	40.316	3.0	9.1	26.2	2.6	2.4	106
444	65.411	54.736	3.3	15.2	15.2	1.6	1.5	123
117	69.603	11.422	5.6	6.6	44.8	4.9	4.4	92

TABLE VII.6

X-Ray Diffraction Data for InP(111) Crystals (a = 0.58690 nm) Using Cu kα Radiation (λ = 0.1540594 nm)

hkl	θ_B (deg)	ϕ (deg)	W_+ (sec)	W_- (sec)	t_{ext} (μm)	t_{abs} (μm)	t_p (μm)	$\lvert F_h \rvert$
111	13.140	0	18.0	18.0	5.9	1.2	1.0	184
222	27.043	0	5.3	5.3	21.6	2.4	2.2	98
224	40.015	19.471	4.3	10.4	16.9	2.7	2.3	151
333	42.999	0.0	4.9	4.9	27.2	3.6	3.2	121
044	47.941	35.264	2.9	12.9	15.6	1.6	1.7	139
135	50.939	28.561	2.9	7.4	26.5	2.9	2.6	104
444	65.411	0	7.0	7.0	34.6	4.8	4.2	123



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Appendix VIII: Resolution of X-Ray Measurements

Consider a normalized Gaussian profile given by

$$I_G(\theta) = \exp\left\{-\left(\frac{\theta-\theta_0}{\sqrt{2}\sigma}\right)^2\right\}, \tag{VIII.1}$$

where θ_0 is the center of the Gaussian peak and σ is the standard deviation parameter, and suppose that a noise signal $\eta(\theta)$ of amplitude n is added, where $n \ll 1$. The resulting signal plus noise is given by

$$I_G(\theta) = \exp\left\{-\left(\frac{\theta-\theta_0}{\sqrt{2}\sigma}\right)^2\right\} + \eta(\theta). \tag{VIII.2}$$

In the presence of this noise, the apparent peak position may shift in either direction by an amount $\Delta\theta$ as a consequence of the added noise, where

$$1-n = \exp\left\{-\left(\frac{\Delta\theta}{\sqrt{2}\sigma}\right)^2\right\}. \tag{VIII.3}$$

Solving for the uncertainty in the peak position results in

$$\Delta\theta = \sigma\sqrt{-2\ln(1-n)}. \tag{VIII.4}$$

For a Gaussian peak the standard deviation σ and the full width at half maximum (FWHM) β are related by

$$\sigma = \frac{\beta}{2\sqrt{2\ln(2)}}, \tag{VIII.5}$$

so that

$$\Delta\theta = \frac{\beta}{2}\sqrt{-\frac{\ln(1-n)}{\ln(2)}}. \tag{VIII.6}$$

For a normalized Gaussian profile with noise amplitude n, the signal-to-noise ratio is given as

$$SNR = \frac{1}{n}, \tag{VIII.7}$$

and thus

$$\Delta\theta = \frac{\beta}{2}\sqrt{-\frac{\ln\left(1-\frac{1}{SNR}\right)}{\ln(2)}}. \tag{VIII.8}$$

Therefore, the uncertainty in the measured peak position is proportional to the peak width,

$$\Delta\theta = C_P\beta, \tag{VIII.9}$$

with constant of proportionality

$$C_P = \frac{1}{2}\sqrt{-\frac{\ln\left(1-\frac{1}{SNR}\right)}{\ln(2)}}. \tag{VIII.10}$$

The emission and detection of x-rays is a random process that is governed by Poisson statistics. For a large number of counts, the Poisson distribution becomes approximately Gaussian with a standard deviation equal to $\sqrt{N_{\text{Eff}}}$, where N_{Eff} is the number of x-ray photons detected. For application to thin films in order to determine the critical layer thickness, in-plane strain, misfit dislocations, and threading dislocations, we estimate the typical range of total counts (intensity in counts per second times the count time in seconds), depending on the type of experimental setup (rotating or fixed anode). In most practical experiments, curve fitting is used in conjunction with an excess of measured points. If there are m points measured within the range of uncertainty for the peak position or peak width, then the effective number of counts is increased to $N_{\text{Eff}} = mN$. Following the usual assumption that the noise amplitude is equal to the standard deviation, the signal-to-noise ratio is

$$SNR = \frac{N_{\text{Eff}}}{\sqrt{N_{\text{Eff}}}} = \sqrt{N_{\text{Eff}}}\ . \tag{VIII.11}$$

Therefore, in terms of the effective counts, Equation VIII.10 is modified to

$$C_P = \frac{1}{2\sqrt{\ln(2)}}\sqrt{-\ln\left(1-\frac{1}{\sqrt{N_{\text{Eff}}}}\right)}. \tag{VIII.12}$$

The additive noise also introduces an error in the determination and measurement of the peak width. For a normalized Gaussian profile, the maximum displacement of the half-intensity point, which is the error in the half width at half maximum (HWHM), is

$$\frac{\Delta\beta}{2} = n\left|\left(\left.\frac{dI_G}{d\theta}\right|_{\theta=\theta_0+\beta/2}\right)^{-1}\right|, \tag{VIII.13}$$

where $\Delta\beta$ is the error in the measurement of the FWHM β. Differentiating the Gaussian profile, we obtain

$$\frac{dI_G}{d\theta} = -\frac{(\theta-\theta_0)\exp\left\{-\left(\frac{\theta-\theta_0}{\sqrt{2}\sigma}\right)^2\right\}}{\sigma^2}, \tag{VIII.14}$$

so that

$$\Delta\beta = \frac{n\beta}{\ln(2)} = \frac{\beta}{SNR\cdot\ln(2)}. \tag{VIII.15}$$

The uncertainty in the peak width is proportional to the peak width,

$$\Delta\beta = C_w\beta, \tag{VIII.16}$$

with a constant of proportionality

$$C_w = \frac{1}{SNR\cdot\ln(2)} = \frac{1}{\sqrt{N_{Eff}}\cdot\ln(2)}. \tag{VIII.17}$$

X-ray profiles from high-quality substrate crystals or pseudomorphic epitaxial layers are sometimes more Lorentzian in character, and it is therefore of interest to consider a normalized Lorentzian profile given by

$$I_L(\theta) = \frac{\left(\frac{\beta}{2}\right)^2}{(\theta-\theta_0)^2+\left(\frac{\beta}{2}\right)^2}, \tag{VIII.18}$$

where θ_0 is the center of the Lorentzian peak and β is the FWHM. Similar to the Gaussian analysis, if a noise signal $\eta(\theta)$ of amplitude n is added, where $n \ll 1$, the resultant profile is

$$I_L(\theta) = \frac{\left(\frac{\beta}{2}\right)^2}{(\theta-\theta_0)^2+\left(\frac{\beta}{2}\right)^2} + \eta(\theta). \tag{VIII.19}$$

The added noise may shift the apparent peak position by an amount $\Delta\theta$, where

$$1-n=\frac{\left(\frac{\beta}{2}\right)^2}{\Delta\theta^2+\left(\frac{\beta}{2}\right)^2}. \tag{VIII.20}$$

Solving, we obtain

$$\Delta\theta=\frac{\beta}{2}\sqrt{\frac{n}{1-n}}\ . \tag{VIII.21}$$

Thus, the uncertainty in the peak position is proportional to the peak width,

$$\Delta\theta=C_P\beta\ , \tag{VIII.22}$$

where the constant of proportionality is

$$C_P=\frac{1}{2}\sqrt{\frac{1}{SNR-1}}=\frac{1}{2}\sqrt{\frac{1}{\sqrt{N_{\mathrm{Eff}}}-1}}\ . \tag{VIII.23}$$

The introduction of noise also affects the determination of the peak width, similar to the case of the Gaussian profile. For a normalized Lorentzian profile, the maximum displacement of the half-intensity point, which is the error in the HWHM, is

$$\frac{\Delta\beta}{2}=n\left|\left(\left.\frac{dI_L}{d\theta}\right|_{\theta=\theta_0+\beta/2}\right)^{-1}\right|. \tag{VIII.24}$$

Differentiating the Lorentzian profile, we obtain

$$\frac{dI_L}{d\theta}=\frac{2\left(\theta-\theta_0\right)\left(\frac{\beta}{2}\right)^2}{\left(\left(\theta-\theta_0\right)^2+\left(\frac{\beta}{2}\right)^2\right)^2}, \tag{VIII.25}$$

so that

$$\Delta\beta=2n\beta=\frac{2\beta}{SNR}\ . \tag{VIII.26}$$

The uncertainty in the peak width is proportional to the peak width,

$$\Delta\beta = C_w \beta \,, \tag{VIII.27}$$

with a constant of proportionality

$$C_w = \frac{2}{SNR} = \frac{2}{\sqrt{N_{\text{Eff}}}} \,. \tag{VIII.28}$$

Figures VIII.1 through VIII.8 explore the dependence of the resolution on the effective number of x-ray counts, the FWHM of the diffraction peak, and the Bragg angle. The resolutions for both the peak position measurements and peak width measurements are considered, as well as both the Gaussian and Lorentzian profile shapes.

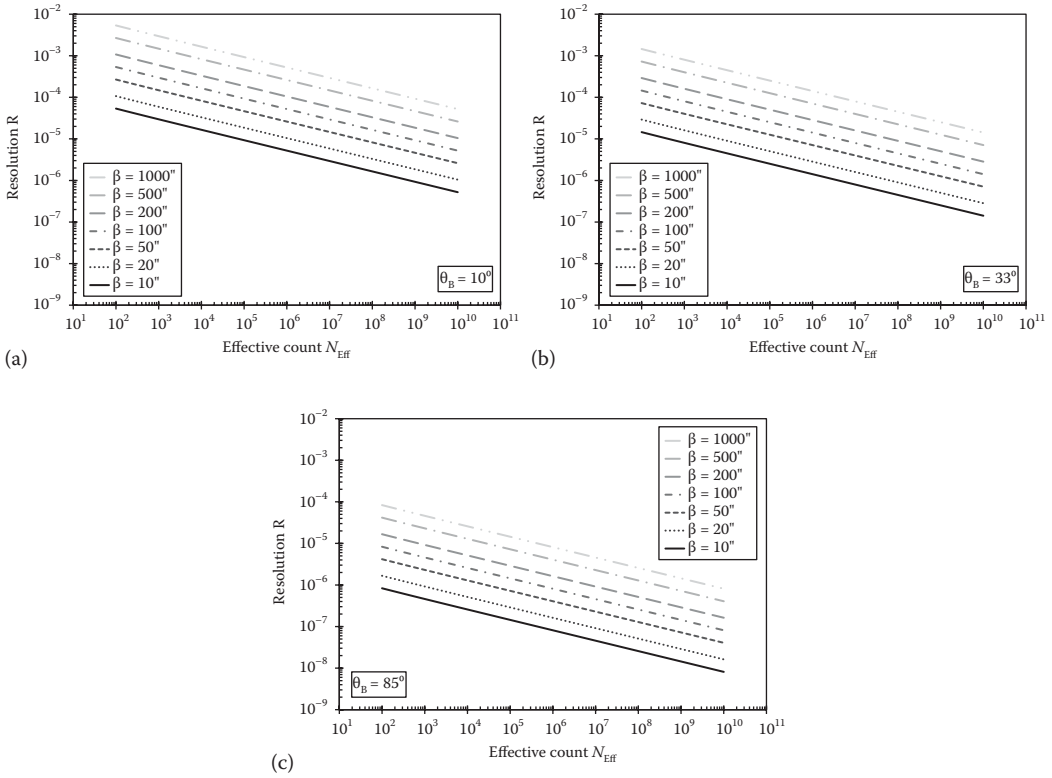


FIGURE VIII.1
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Gaussian in character, as a function of the effective number of counts and with peak FWHM as a parameter, for Bragg angles of (a) 10°, (b) 33°, and (c) 85°. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

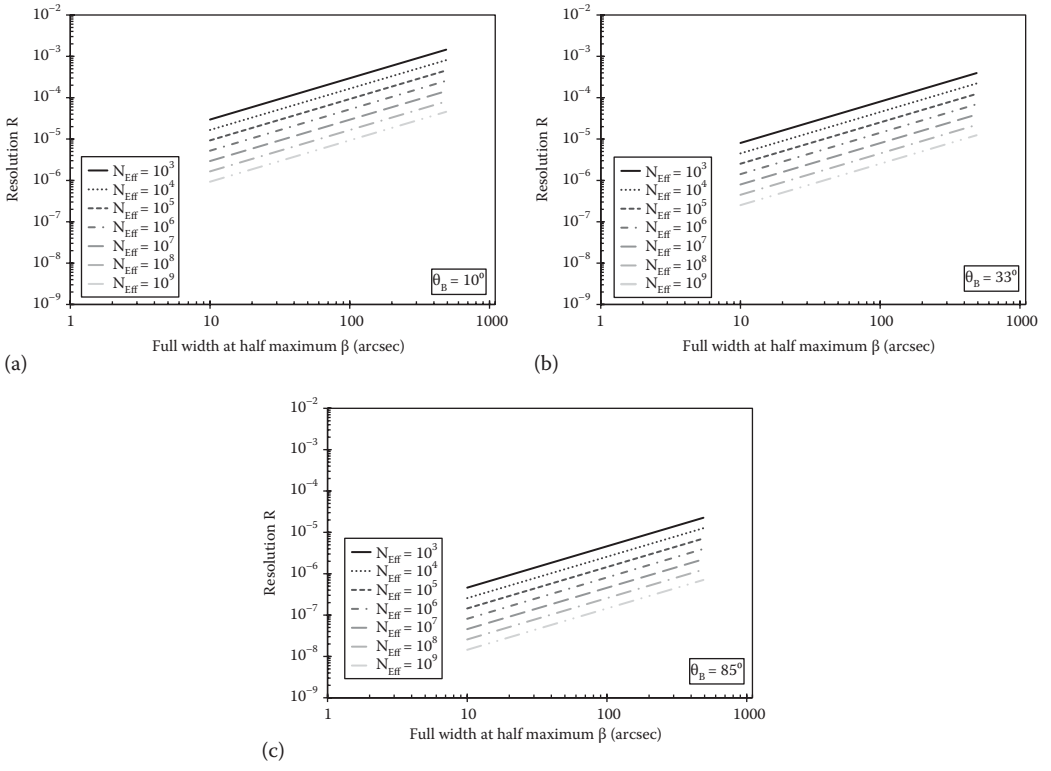


FIGURE VIII.2
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Gaussian in character, as a function of the peak FWHM and with the effective number of counts as a parameter, for Bragg angles of (a) 10°, (b) 33°, and (c) 85°. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

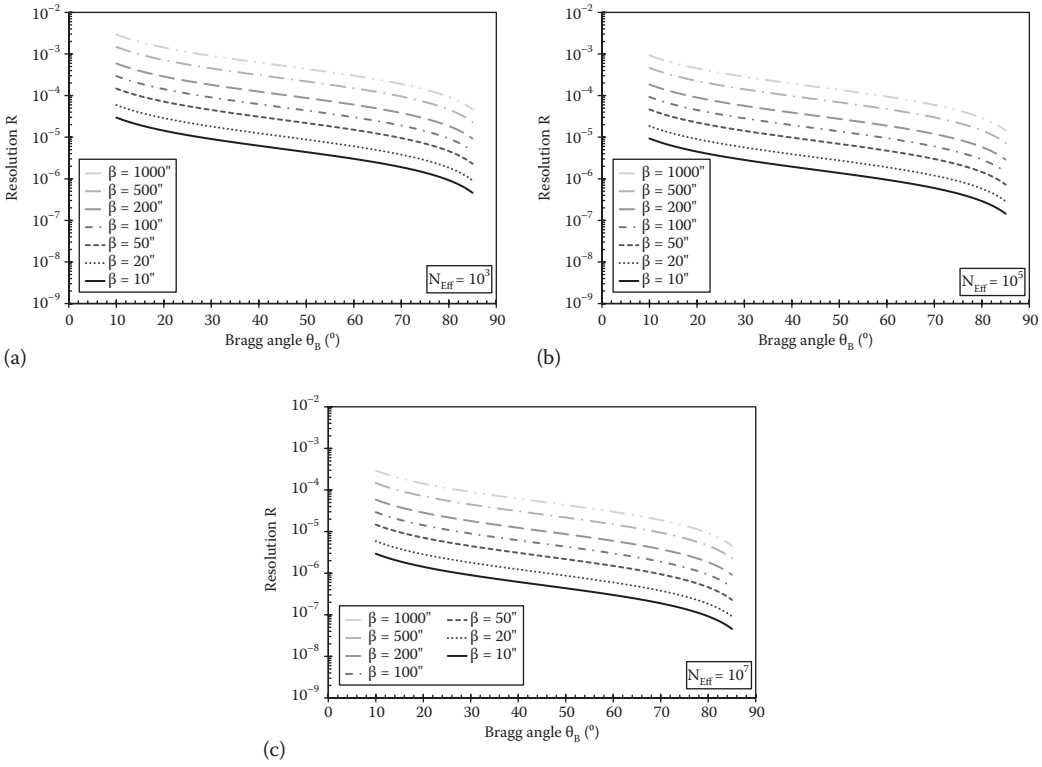


FIGURE VIII.3
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Gaussian in character, as a function of the Bragg angle and with peak FWHM as a parameter, for an effective number of counts equal to (a) 10^3, (b) 10^5, and (c) 10^7. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

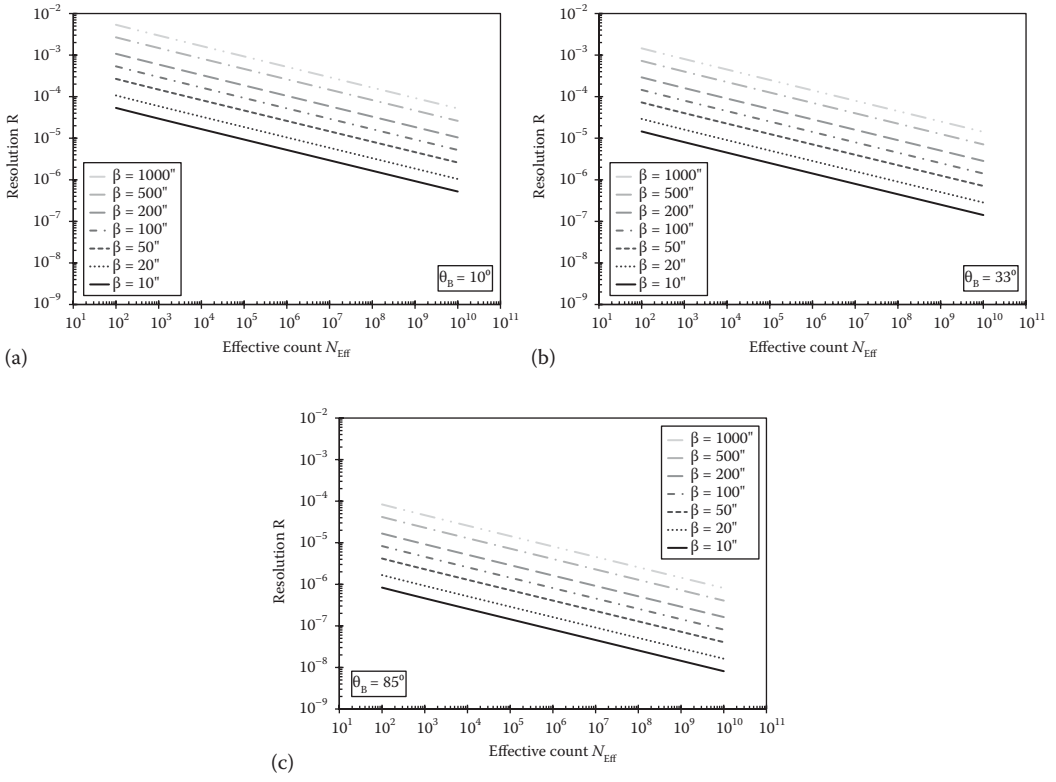


FIGURE VIII.4
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Lorentzian in character, as a function of the effective number of counts and with peak FWHM as a parameter, for Bragg angles of (a) 10°, (b) 33°, and (c) 85°. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

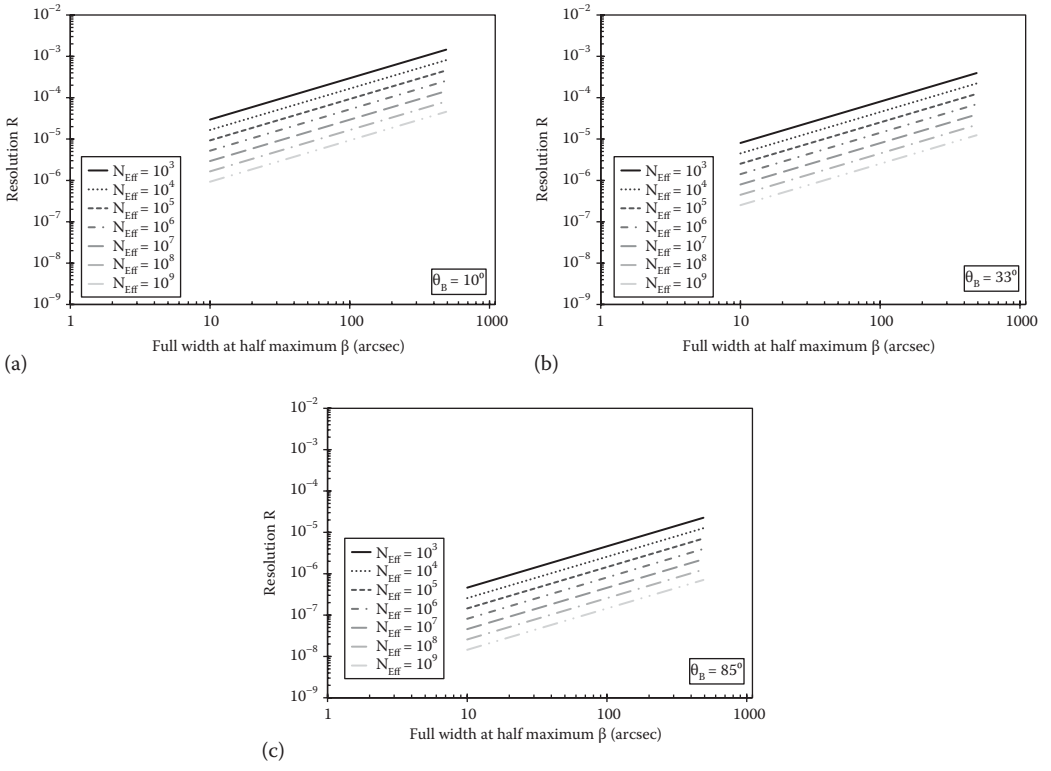


FIGURE VIII.5
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Lorentzian in character, as a function of the peak FWHM and with the effective number of counts as a parameter, for Bragg angles of (a) 10°, (b) 33°, and (c) 85°. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

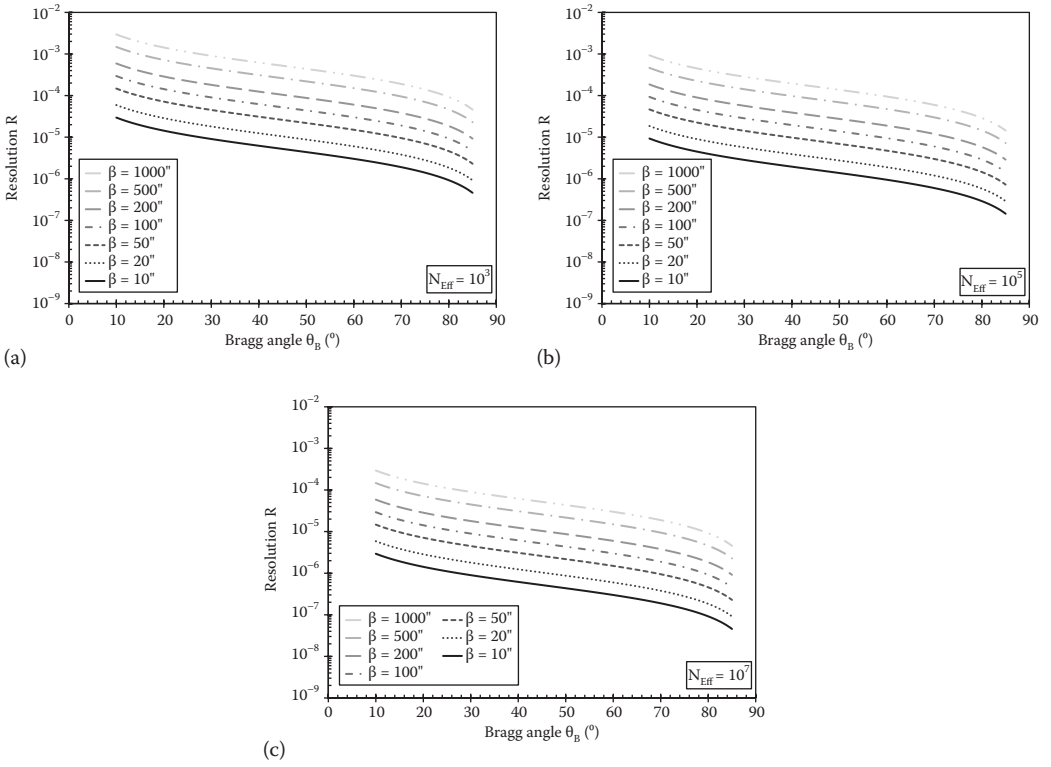


FIGURE VIII.6
Absolute resolution of peak position determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Lorentzian in character, as a function of the Bragg angle and with peak FWHM as a parameter, for an effective number of counts equal to (a) 10^3, (b) 10^5, and (c) 10^7. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

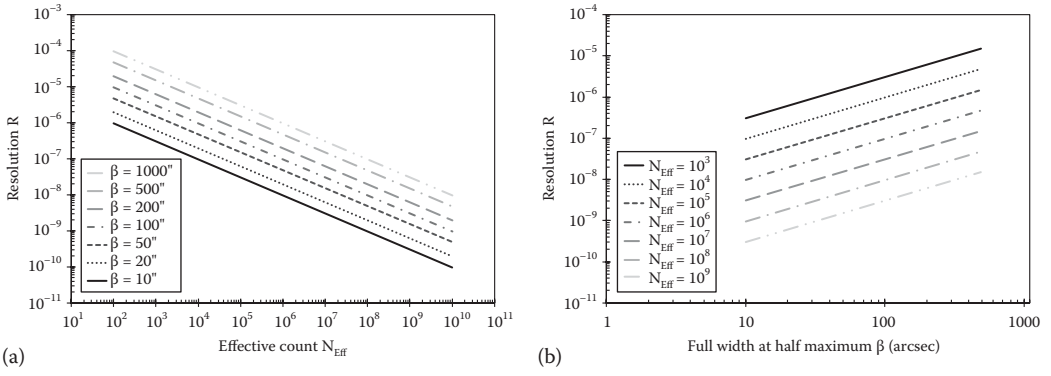


FIGURE VIII.7
Absolute resolution of peak width determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Gaussian in character, (a) as a function of the number of effective counts and with peak FWHM as a parameter, and (b) as a function of peak FWHM and with the effective number of counts as a parameter. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)

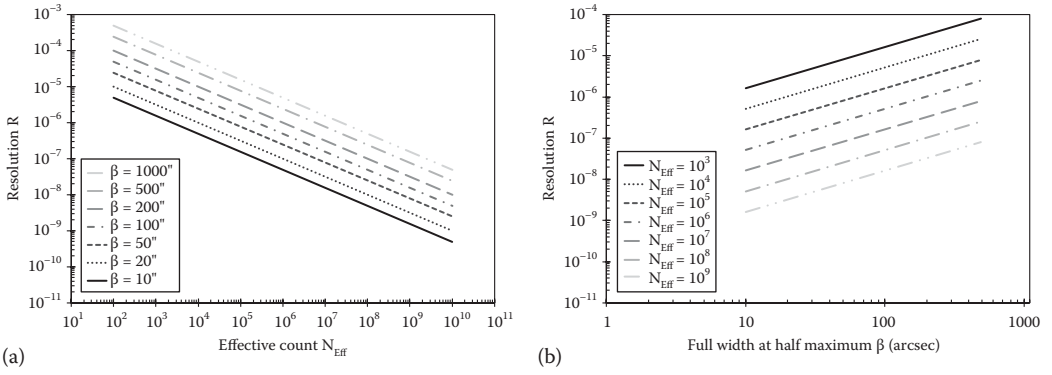


FIGURE VIII.8
Absolute resolution of peak width determination by an x-ray diffraction experiment, with a diffraction profile that is assumed to be Lorentzian in character, (a) as a function of the number of effective counts and with peak FWHM as a parameter, and (b) as a function of peak FWHM and with the effective number of counts as a parameter. (Reprinted from T. Kujofsa and J. E. Ayers, *Intl. J. High Speed Electron. Syst.*, 24, 15500007 [2015]. With permission. Copyright 2015, World Scientific.)



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com