

Integrated Optics

Sixth Edition

Robert G. Hunsperger

Integrated Optics

Theory and Technology

Sixth Edition



Robert G. Hunsperger
Department of Electrical Engineering
University of Delaware
140 Evans Hall
Newark, DE 19716–3130, USA
hunsperg@eecis.udel.edu

ISBN 978-0-387-89774-5

e-ISBN 978-0-387-89775-2

DOI 10.1007/b98730

Library of Congress Control Number: 2008942780

© Springer Science+Business Media, LLC 1982, 1984, 1991, 1995, 2002, 2009

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

springer.com

To our grandchildren,
Cameron Hunsperger Wilson
and
Caitlin Margaret Wilson

Preface to the Sixth Edition

The field of integrated optics has changed considerably in the past 7 years. The dimensions of photonic devices to be integrated have decreased by several orders of magnitude. Previously, the topics covered in chapters of this book might be described as belonging to the field of microphotonics, since they involved photons of light interacting with physical structures having dimensions mostly on the order of micrometers. The two exceptions were gratings and quantum wells, some of which had periodicity on the order of 100 nm. As fabrication techniques have progressed it has become possible to produce nanometer-sized structures such as quantum wires, quantum dots, holographic optical elements (HOEs), and photonic crystals (PhCs). These new devices have greatly changed both the size and performance of integrated-optic circuits.

In response to these new developments, a new chapter has been added to this sixth edition: Nanophotonics. The important topics of confinement of photons and electrons, photonic crystals and nanophotonic devices are covered in this new chapter. Techniques for fabricating and evaluating nanostructures are also described. All the other chapters have also been updated to include new developments and literature references. Additional practice problems have been added to all chapters, and an updated booklet of problem solutions is available.

Newark, DE
March 2009

R.G. Hunsperger

Preface to the Fifth Edition

Since the publication of the fourth edition a number of new areas in the general field of integrated optics have gained prominence. The telecommunications industry has continued the implementation of fiber-optic networks, not only for trunk lines but also to bring optical fiber waveguides to the office and home. This widespread use of optical fibers has created a demand for effective, but relatively inexpensive, optical amplifiers, couplers, and switches. This demand has been met by erbium-doped fiber amplifiers (EDFAs) and by couplers and switches fabricated in glasses and in polymer materials. New systems have been developed that incorporate micromachined mechanical elements along with optical and electronic devices in an integrated structure. These are called micro-optical-electro-mechanical systems (MOEMs or MEMs). Most recently, optoelectronic devices and optical integrated circuits have found use in wireless systems, in which optical fibers are used to carry signals over long distances while RF and microwave transceivers complete the final link to the user.

In response to these new developments, four new chapters have been added to this fifth edition: Polymer and Fiber Integrated Optics, Optical Amplifiers, Micro-electro-optical-mechanical Devices (MEMs), and Optoelectronic Devices in Wireless Systems. All the other chapters have also been updated to include new developments and literature references. The new chapters include additional practice problems, and an updated booklet of problem solutions is available.

The author wishes to thank Dr. Wei Liu for his help in reading the page proofs and identifying necessary revisions.

Newark, DE
March 2002

R.G. Hunsperger

Preface to the Fourth Edition

Once again it has become necessary to produce a new edition in order to update material provided in earlier editions and to add new descriptions of recently emerging technology. All of the chapters have been revised to include new developments, and to incorporate additional literature references.

In the past few years there has been a vast expansion of worldwide telecommunications and data transmission networks. In many localities fiber-to-the-home and integrated services digital networks (ISDN) have become a reality. Many people are now logging-on to the Internet and the World Wide Web. The growth of these networks has created a strong demand for inexpensive, yet efficient and reliable, integrated optic components such as signal splitters, couplers and multiplexers. Because of this demand, there has been a great deal of work recently on devices made using polymers and glasses. Descriptions of these components have been added to the book in the appropriate chapters.

A number of new practice problems have been added, and an updated booklet of problem solutions is available. The supplementary series of videotaped lectures described in the preface to earlier editions continues to be available. Inquiries regarding these materials should be sent directly to the author.

The author wishes to thank Mrs. Barbara Westog, who helped with the organization of new material and typed the revisions.

Newark, July 1995

R.G. Hunsperger

Preface to the Third Edition

The field of integrated optics is continuing to develop at a rapid pace, necessitating the writing of this third edition in order to update the material contained in earlier editions. All of the chapters have been revised to reflect the latest developments in the field and a new chapter has been added to explain the important topic of newly invented quantum well devices. These promise to significantly improve the operating characteristics of lasers, modulators and detectors.

The trend of telecommunications toward the use of single mode systems operating at the longer wavelengths of 1.3 and 1.55 μm has been explained and documented with illustrations of recently developed devices and systems. In this regard, broader coverage of GaInAsP devices and optical integrated circuits has been provided, and the new growth techniques of molecular beam epitaxy (MBE) and metal-organic chemical vapor deposition (MOCVD) have been described. The extensive development of hybrid optical integrated circuits in lithium niobate has also been described. Notably, this progress has led to the production of the first commercially available optical integrated circuits.

A number of new practice problems have been added. An updated booklet of problem solution is available, and the supplementary series of videotaped lectures described in the preface to the first edition has been expanded and updated. Inquiries regarding these materials should be sent directly to the author.

The author wishes to thank Mr. Garfield Simms, who has generated the artwork for a number of the new illustrations which appear in this edition, and Mrs. Barbara Westog, who typed the revisions.

Newark, January 1991

R.G. Hunsperger

Preface to the Second Edition

Our intent in producing this book was to provide a text that would be comprehensive enough for an introductory course in integrated optics, yet concise enough in its mathematical derivations to be easily readable by a practicing engineer who desires an overview of the field. The response to the first edition has indeed been gratifying; unusually strong demand has caused it to be sold out during the initial year of publication, thus providing us with an early opportunity to produce this updated and improved second edition.

This development is fortunate, because integrated optics is a very rapidly progressing field, with significant new research being regularly reported. Hence, a new chapter (Chap. 17) has been added to review recent progress and to provide numerous additional references to the relevant technical literature. Also, thirty-five new problems for practice have been included to supplement those at the ends of chapters in the first edition. Chapters 1 through 16 are essentially unchanged, except for brief updating revisions and corrections of typographical errors.

Because of the time limitations imposed by the need to provide an uninterrupted supply of this book to those using it as a course text, it has been possible to include new references and to briefly describe recent developments only in Chapter 17. However, we hope to provide details of this continuing progress in a future edition.

The author wishes to thank Mr. Mark Bendett, Mr. Jung-Ho Park, and Dr. John Zavada for their valuable help in locating typographical errors and in developing new problems for this edition.

Newark, December 1983

R.G. Hunsperger

Preface to the First Edition

This book is an introduction to the theory and technology of integrated optics for graduate students in electrical engineering, and for practicing engineers and scientists who wish to improve their understanding of the principles and applications of this relatively new, and rapidly growing field.

Integrated Optics is the name given to a new generation of opto-electronic systems in which the familiar wires and cables are replaced by light-waveguiding optical fibers, and conventional integrated circuits are replaced by optical integrated circuits (OIC's). In an OIC, the signal is carried by means of a beam of light rather than by an electrical current, and the various circuit elements are interconnected on the substrate wafer by optical waveguides. Some advantages of an integrated-optic system are reduced weight, increased bandwidth (or multiplexing capability), resistance to electromagnetic interference, and low loss signal transmission.

Because of the voluminous work that has been done in the field of integrated optics since its inception in the late 1960's, the areas of fiber optics and optical integrated circuits have usually been treated separately at conferences and in textbooks. In the author's opinion, this separation is unfortunate because the two areas are closely related. Nevertheless, it cannot be denied that it may be a practical necessity. Hence, this book includes an overview of the entire field of integrated optics in the first chapter, which relates the work on optical integrated circuits to progress in fiber-optics development. Specific examples of applications of both fibers and the OIC's are given in the final chapter. The remaining chapters of the book are devoted to the detailed study of the phenomena, devices and technology of optical integrated circuits.

This book is an outgrowth of a graduate level, single-semester course in integrated optics taught first at the University of Southern California in 1975 and later at the University of Delaware. The course has also been produced as a series of 20 color videotaped lectures, which can be used along with this book for self-study of the subject. A booklet of solutions to the problems given at the end of the chapters is also available. Inquiries regarding these supplementary materials should be sent directly to the author.

The author wishes to thank those persons who have contributed to making this book a reality. In particular, the critical comments and constructive suggestions provided by Dr. T. Tamir throughout the preparation of the manuscript have been

most helpful. The continuing support and encouragement of Dr. H. Lotsch are also greatly appreciated. The competent and efficient typing of the manuscript by Mrs. Anne Seibel and Miss Jacqueline Gregg has greatly facilitated timely publication.

Newark, April 1982

R.G. Hunsperger

Contents

1	Introduction	1
1.1	Advantages of Integrated Optics	2
1.1.1	Comparison of Optical Fibers with Other Interconnectors	3
1.1.2	Comparison of Optical Integrated Circuits with Electrical Integrated Circuits	7
1.2	Substrate Materials for Optical Integrated Circuits	8
1.2.1	Hybrid Versus Monolithic Approach	9
1.2.2	III-V and II-VI Ternary Systems	10
1.2.3	Hybrid OIC's in LiNbO ₃	11
1.2.4	Organization of this Book	12
Problems		15
References		15
2	Optical Waveguide Modes	17
2.1	Modes in a Planar Waveguide Structure	17
2.1.1	Theoretical Description of the Modes of a Three-Layer Planar Waveguide	17
2.1.2	Cutoff Conditions	20
2.1.3	Experimental Observation of Waveguide Modes	21
2.2	The Ray-Optic Approach to Optical Mode Theory	25
2.2.1	Ray Patterns in the Three-Layer Planar Waveguide	26
2.2.2	The Discrete Nature of the Propagation Constant β	28
Problems		30
References		31
3	Theory of Optical Waveguides	33
3.1	Planar Waveguides	33
3.1.1	The Basic Three-Layer Planar Waveguide	33
3.1.2	The Symmetric Waveguide	36
3.1.3	The Asymmetric Waveguide	38

3.2	Rectangular Waveguides	39
3.2.1	Channel Waveguides	40
3.2.2	Strip-Loaded Waveguides	45
Problems		48
References		49
4	Waveguide Fabrication Techniques	53
4.1	Deposited Thin Films	53
4.1.1	Sputtered Dielectric Films	53
4.1.2	Deposition from Solutions	56
4.1.3	Organosilicon Films	57
4.2	Substitutional Dopant Atoms	57
4.2.1	Diffused Dopants	57
4.2.2	Ion Exchange and Migration	59
4.2.3	Ion Implantation	60
4.3	Carrier-Concentration-Reduction Waveguides	62
4.3.1	Basic Properties of Carrier-Concentration-Reduction Waveguides	62
4.3.2	Carrier Removal by Proton Bombardment	64
4.4	Epitaxial Growth	65
4.4.1	Basic Properties of Epitaxially Grown Waveguides	65
4.4.2	Ga($1-x$)Al _x As Epitaxially Grown Waveguides	66
4.4.3	Epitaxial Waveguides in Other III-V and II-VI and IV Materials	70
4.4.4	Molecular Beam Epitaxy	71
4.4.5	Metal-Organic Chemical Vapor Deposition	72
4.5	Electro-Optic Waveguides	73
4.6	Oxidation	74
4.7	Methods for Fabricating Channel Waveguides	75
4.7.1	Ridged Waveguides Formed by Etching	75
4.7.2	Strip-Loaded Waveguides	77
4.7.3	Masked Ion Implantation, Diffusion or Ion Exchange	77
4.7.4	Focused Beam Writing Techniques	78
Problems		79
References		81
5	Polymer and Fiber Integrated Optics	85
5.1	Types of Polymers	85
5.2	Polymer Processing	87
5.2.1	Processing of Polystyrene	87
5.2.2	Processing of Polyimide	89
5.2.3	Post-Deposition Processing	89
5.3	Applications of Polymer Waveguide Interconnections	90
5.4	Polymer Waveguide Devices	93
5.4.1	Passive Polymer Devices	93
5.4.2	Active Polymer Devices	96

5.5	Optical Fiber Waveguide Devices	98
Problems		103
References		104
6	Losses in Optical Waveguides	107
6.1	Scattering Losses	107
6.1.1	Surface Scattering Loss	108
6.2	Absorption Losses	110
6.2.1	Interband Absorption	110
6.2.2	Free Carrier Absorption	112
6.3	Radiation Losses	116
6.3.1	Radiation Loss from Planar and Straight Channel Waveguides	117
6.3.2	Radiation Loss from Curved Channel Waveguides	117
6.4	Measurement of Waveguide Losses	120
6.4.1	End-Fire Coupling to Waveguides of Different Length	120
6.4.2	Prism-Coupled Loss Measurements	122
6.4.3	Scattering Loss Measurements	123
Problems		125
References		127
7	Waveguide Input and Output Couplers	129
7.1	Fundamentals of Optical Coupling	129
7.2	Transverse Couplers	130
7.2.1	Direct Focusing	130
7.2.2	End-Butt Coupling	131
7.3	Prism Couplers	135
7.4	Grating Couplers	139
7.4.1	Basic Theory of the Grating Coupler	139
7.4.2	Grating Fabrication	141
7.5	Tapered Couplers	143
7.6	Tapered Mode Size Converters	144
7.7	Fiber to Waveguide Couplers	145
7.7.1	Butt Coupling	145
7.7.2	High Density Multifiber Connectors	148
Problems		149
References		151
8	Coupling Between Waveguides	153
8.1	Multilayer Planar Waveguide Couplers	153
8.2	Dual-Channel Directional Couplers	154
8.2.1	Operating Characteristics of the Dual-Channel Coupler	155
8.2.2	Coupled-Mode Theory of Synchronous Coupling	157

8.2.3	Methods of Fabricating Dual-Channel Directional Coupplers	160
8.2.4	Applications Involving Directional Couplers	164
8.3	Butt-Coupled Ridge Waveguides	164
8.4	Branching Waveguide Couplers	164
8.5	Optical Fiber Couplers and Splitters	166
	Problems	167
	References	168
9	Electro-Optic Modulators	171
9.1	Basic Operating Characteristics of Switches and Modulators	171
9.1.1	Modulation Depth	171
9.1.2	Bandwidth	172
9.1.3	Insertion Loss	172
9.1.4	Power Consumption	173
9.1.5	Isolation	173
9.2	The Electro-Optic Effect	174
9.3	Single-Waveguide Electro-Optic Modulators	175
9.3.1	Phase Modulation	175
9.3.2	Polarization Modulation	177
9.3.3	Intensity Modulation	177
9.3.4	Electro-Absorption Modulation	178
9.4	Dual-Channel Waveguide Electro-Optic Modulators	181
9.4.1	Theory of Operation	181
9.4.2	Operating Characteristics of Dual-Channel Modulators ..	183
9.5	Mach-Zehnder Type Electro-Optic Modulators	187
9.6	Electro-Optic Modulators Employing Reflection or Diffraction ..	188
9.6.1	Bragg-Effect Electro-Optic Modulators	188
9.6.2	Electro-Optic Reflection Modulators	190
9.7	Comparison of Waveguide Modulators to Bulk Electro-Optic Modulators	191
9.8	Traveling Wave Electrode Configurations	193
	Problems	195
	References	198
10	Acousto-Optic Modulators	201
10.1	Fundamental Principles of the Acousto-Optic Effect	201
10.2	Raman-Nath-Type Modulators	203
10.3	Bragg-Type Modulators	204
10.4	Bragg-Type Beam Deflectors and Switches	208
10.5	Performance Characteristics of Acoustic-Optic Modulators and Beam Deflectors	210
10.6	Accusto-Optic Frequency Shifters	214
	Problems	217
	References	219

11 Basic Principles of Light Emission in Semiconductors	221
11.1 A Microscopic Model for Light Generation and Absorption in a Crystalline Solid	221
11.1.1 Basic Definitions	221
11.1.2 Conservation of Energy and Momentum	224
11.2 Light Emission in Semiconductors	226
11.2.1 Spontaneous Emission	226
11.2.2 Stimulated Emission	232
11.3 Lasing	234
11.3.1 Semiconductor Laser Structures	235
11.3.2 Lasing Threshold	235
11.3.3 Efficiency of Light Emission	237
Problems	238
References	239
12 Semiconductor Lasers	241
12.1 The Laser Diode	241
12.1.1 Basic Structure	241
12.1.2 Optical Modes	242
12.1.3 Lasing Threshold Conditions	243
12.1.4 Output Power and Efficiency	248
12.2 The Tunnel-Injection Laser	250
12.2.1 Basic Structure	250
12.2.2 Lasing Threshold Conditions	252
12.3 Polymer Lasers	252
12.4 New Semiconductor Materials for New Wavelengths	253
12.4.1 Gallium Nitride Lasers	253
12.4.2 Silicon Lasers	254
Problems	255
References	257
Supplementary Reading on Semiconductor-Laser Fundamentals	258
13 Optical Amplifiers	259
13.1 Optical Fiber Amplifiers	259
13.1.1 Erbium Doped Fiber Amplifiers	260
13.1.2 Raman Optical Fiber Amplifiers	263
13.1.3 Other Optical Fiber Amplifiers	264
13.2 Non-Fiber Ion-Doped Optical Amplifiers	265
13.3 Semiconductor Optical Amplifiers	265
13.3.1 Integrated Semiconductor Optical Amplifiers	268
13.4 Comparison of Ion-Doped Fiber Amplifiers with SOAs	269
13.4.1 Wavelength Range	269
13.4.2 Performance Characteristics	269
13.5 Gain Equalization	271
13.6 Fiber Lasers	271

Problems	273
References	274
Supplementary Reading on Optical Amplifiers	275
14 Heterostructure, Confined-Field Lasers	277
14.1 Basic Heterojunction Laser Structures	278
14.1.1 Single Heterojunction (SH) Lasers	278
14.1.2 Double Heterostructure (DH) Lasers	279
14.2 Performance Characteristics of the Heterojunction Laser	280
14.2.1 Optical Field Confinement	280
14.2.2 Carrier Confinement	283
14.2.3 Comparison of Laser Emission Characteristics	284
14.3 Control of Emitted Wavelength	285
14.3.1 $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ Lasers for Fiber-Optic Applications	285
14.3.2 Lasers Made of Quaternary Materials	287
14.3.3 Long-Wavelength Lasers	287
14.4 Advanced Heterojunction Laser Structures	288
14.4.1 Stripe Geometry Lasers	288
14.4.2 Single-Mode Lasers	288
14.4.3 Integrated Laser Structures	291
14.5 Reliability	295
14.5.1 Catastrophic Failure	295
14.5.2 Gradual Degradation	296
14.6 Vertical Cavity Lasers	296
Problems	298
References	299
Supplementary Reading on Heterojunction Lasers	301
15 Distributed-Feedback Lasers	303
15.1 Theoretical Considerations	303
15.1.1 Wavelength Dependence of Bragg Reflections	303
15.1.2 Coupling Efficiency	305
15.1.3 Lasing with Distributed Feedback	308
15.2 Fabrication Techniques	309
15.2.1 Effects of Lattice Damage	310
15.2.2 Grating Location	310
15.2.3 DBR Lasers	313
15.3 Performance Characteristics	315
15.3.1 Wavelength Selectability	315
15.3.2 Optical Emission Linewidth	317
15.3.3 Stability	317
15.3.4 Commercially Available DFB Lasers	319
15.4 Nanoscale DFB Lasers	319
15.4.1 Semiconductor Air Bragg Reflector Lasers	320
15.4.2 Quantum Dot DFB Lasers	321

Problems	321
References	322
16 Direct Modulation of Semiconductor Lasers	325
16.1 Basic Principles of Direct Modulation	325
16.1.1 Amplitude Modulation	325
16.1.2 Pulse Modulation	328
16.1.3 Frequency Modulation	330
16.2 Microwave Frequency Modulation of Laser Diodes	331
16.2.1 Summary of Early Experimental Results	332
16.2.2 Factors Limiting Modulation Frequency	332
16.2.3 Design of Laser Diode Packages for Microwave Modulation	336
16.3 Monolithically Integrated Direct Modulators	337
16.4 Amplified Laser Modulation	339
16.5 Direct Modulation of Quantum Dot Lasers	339
16.6 Future Prospects for Microwave Modulation of Laser Diodes	340
Problems	340
References	342
Supplementary Reading on Modulation of Laser Diodes	344
17 Integrated Optical Detectors	345
17.1 Depletion Layer Photodiodes	345
17.1.1 Conventional Discrete Photodiodes	345
17.1.2 Waveguide Photodiodes	348
17.1.3 Effects of Scattering and Free-Carrier Absorption	349
17.2 Specialized Photodiode Structures	350
17.2.1 Schottky-Barrier Photodiode	351
17.2.2 Avalanche Photodiodes	351
17.2.3 p-i-n Photodiodes	353
17.2.4 Metal-Semiconductor-Metal Photodiodes	354
17.3 Techniques for Modifying Spectral Response	355
17.3.1 Hybrid Structures	355
17.3.2 Heteroepitaxial Growth	356
17.3.3 Proton Bombardment	360
17.3.4 Electro-Absorption	363
17.4 Factors Limiting Performance of Integrated Detectors	366
17.4.1 High Frequency Cutoff	366
17.4.2 Linearity	367
17.4.3 Noise	367
Problems	368
References	371

18 Quantum-Well Devices	375
18.1 Quantum Wells and Superlattices	375
18.2 Quantum-Well Lasers	377
18.2.1 Single-Quantum-Well Lasers	377
18.2.2 Multiple Quantum Well Lasers	380
18.3 Quantum-Well Modulators and Switches	384
18.3.1 Electro-Absorption Modulators	384
18.3.2 Electro-Optic Effect in Quantum Wells	388
18.3.3 Multiple Quantum Well Switches	390
18.4 Quantum-Well Detectors	392
18.4.1 Photoconductive Detectors	392
18.4.2 MQW Avalanche Photodiodes	392
18.5 Self-Electro-Optic Effect Devices	393
18.6 Quantum-Well Devices in OEIC's	394
18.6.1 Integrated Laser/Modulators	395
18.6.2 A Four-Channel Transmitter Array with MQW Lasers	396
Problems	398
References	399
Supplementary Reading on Quantum Wells	401
19 Micro-Optical-Electro-Mechanical Devices	403
19.1 Basic Equations of Mechanics	404
19.1.1 Axial Stress and Strain	404
19.1.2 Thin Membranes	405
19.1.3 Cantilever Beams	406
19.1.4 Torsion Plates	407
19.2 Thin Membrane Devices	408
19.3 Cantilever Beam Devices	411
19.4 Torsional Devices	413
19.5 Optical Elements	417
19.6 Future Directions in MOEMS Development	418
19.7 Mechanical Properties of Silicon	419
Problems	419
References	420
20 Applications of Integrated Optics and Current Trends	423
20.1 Applications of Optical Integrated Circuits	423
20.1.1 RF Spectrum Analyzer	423
20.1.2 Monolithic Wavelength-Multiplexed Optical Source	426
20.1.3 Analog-to-Digital Converter (ADC)	428
20.1.4 Integrated-Optic Doppler Velocimeter	429
20.1.5 An IO Optical Disk Readhead	430
20.1.6 OIC Temperature Sensor	432
20.1.7 IO High Voltage Sensor	433

20.1.8	IO Wavelength Meters and Spectrum Analyzers	434
20.1.9	IO Chemical Sensors	435
20.2	Opto-Electronic Integrated Circuits	436
20.2.1	An OEIC Transmitter	436
20.2.2	An OEIC Receiver	437
20.2.3	An OEIC Phased-Array Antenna Driver	438
20.3	Devices and Systems for Telecommunications	439
20.3.1	Trends in Optical Telecommunications	439
20.3.2	New Devices for Telecommunications	444
Problems	447
References	447
21	Photonic and Microwave Wireless Systems	451
21.1	Merging of Photonics and Microwave Technology	451
21.2	Fiber-Optic Transmission of RF and Microwave Signals	453
21.2.1	Basic Principles	454
21.2.2	Device Performance	456
21.2.3	System Performance	458
21.3	Microwave Carrier Generation by Optical Techniques	459
21.4	Future Projections	463
Problems	464
References	465
22	Nanophotonics	469
22.1	Dimensions	469
22.2	Properties of Electrons and Photons	469
22.3	Confinement of Photons and Electrons	471
22.4	Photonic Crystals	472
22.4.1	Classes of Photonic Crystals	472
22.4.2	Comparison of Electrons in Semiconductor Crystals to Photons in Photonic Crystals	473
22.5	Fabrication of Nanostructures	477
22.5.1	Molecular Beam Epitaxy	478
22.5.2	Metalorganic Vapor Phase Epitaxy	478
22.5.3	Nanoscale Lithography	479
22.5.4	Nanomachining	481
22.6	Characterization and Evaluation of Nanostructures	485
22.6.1	Available Tools	485
22.6.2	Scanning Electron Microscope	485
22.6.3	Reflection High-Energy Electron Diffraction	486
22.7	Nanophotonic Devices	487
22.7.1	Waveguides	487
22.7.2	Couplers	491
22.7.3	Resonators	493
22.7.4	Light Emitters	495

22.7.5 Photodetectors	496
22.7.6 Sensors	497
22.8 Future Projections for Integrated Optics and Nanophotonics	499
Problems	501
References	501
Index	507

Chapter 1

Introduction

The transmission and processing of signals carried by optical beams rather than by electrical currents or radio waves has been a topic of great interest ever since the early 1960s, when the development of the laser first provided a stable source of coherent light for such applications. Laser beams can be transmitted through the air, but atmospheric variations cause undesirable changes in the optical characteristics of the path from day to day, and even from instant to instant. Laser beams also can be manipulated for signal processing, but that requires optical components such as prisms, lenses, mirrors, electro-optic modulators and detectors. All of this equipment would typically occupy a laboratory bench tens of feet on a side, which must be suspended on a vibration-proof mount. Such a system is tolerable for laboratory experiments, but is not very useful in practical applications. Thus, in the late 1960s, the concept of “integrated optics” emerged, in which wires and radio links are replaced by light-waveguiding optical fibers rather than by through-the-air optical paths, and conventional electrical integrated circuits are replaced by miniaturized optical integrated circuits (OIC’s), also known as photonic integrated circuits (PIC’s)

For a historical overview of the first years of integrated optics, the reader is referred to the books edited by Tamir [1] and Miller et al. [2]. During the later years of the 1970s, several factors combined to bring integrated optics out of the laboratory and into the realm of practical application; these were the development of low-loss optical fibers and connectors, the creation of reliable CW GaAlAs and GaInAsP laser diodes, and the realization of photolithographic microfabrication techniques capable of submicron linewidths. In the 1980s, optical fibers largely replaced metallic wires in telecommunications, and a number of manufacturers began production of optical integrated circuits for use in a variety of applications. In the 1990s, the incorporation of optical fibers into telecommunications and data-transmission networks has been extended to the subscriber loop in many systems. This provides an enormous bandwidth for multichannel transmission of voice, video and data signals. Access to worldwide communications and data banks has been provided by computer networks such as the Internet. We are in the process of developing what some have called the “Information superhighway.” The implementation of this technology has provided continuing impetus to the development of new integrated optic devices and systems into the beginning years of the 21st century.

Another technological advance that has encouraged the development of new integrated optic devices in recent years is the availability of improved fabrication methods. Microtechnology, which involves dimensions on the order of micrometers, has evolved into nanotechnology, in which nanometer-sized features are routinely produced. This new area of nanophotonics, which includes the fabrication of photonic crystals, is described in Chapter 22.

Because of the very broad scope of the field of integrated optics, it is common practice to consider optical fiber waveguides and optical integrated circuits as two separate areas of study, even though they are closely related. For example, the first conference on integrated optics, *per se*, sponsored by the Optical Society of America, in February 1972 [3] included sessions on both optical fibers and on optical integrated circuits. However, by February 1974, when the second conference of this series [4] was held, only two papers were presented on optical fibers [5, 6]. Subsequent conferences in this biennial series have included only papers on OIC's. In this book, we will concentrate mainly on optical integrated circuits, but we consider first the advantages of a combined fiber-optic OIC system in order to put the subject matter in proper perspective.

The integrated optics approach to signal transmission and processing offers significant advantages in both performance and cost when compared to conventional electrical methods. In this chapter, we consider those advantages in order to generate an understanding of the motivating force behind the development of integrated optics. We also examine the basic question of which substrate materials are most advantageous for the fabrication of OIC's, and whether a hybrid or a monolithic approach should be used.

1.1 Advantages of Integrated Optics

To consider the advantages of a fiber-optic OIC system as compared to its electrical counterpart, we show in Fig. 1.1 a hypothetical fiber-optic OIC system for optical communications that can be used to illustrate many of the special advantages of the integrated optic approach. In this system, the transmitter and receiver are each contained on an OIC chip, and the two are interconnected by means of an optical fiber waveguide. The elemental devices of the system will be explained in detail in later chapters, but, for now, let us consider only their general functions. The light sources are integrated laser diodes of the distributed feedback (DFB) type, emitting at different wavelengths λ_1 , and λ_2 . Only two diodes are shown for simplicity, but perhaps hundreds would be used in a practical system. Since the light emitted by each laser is at a different wavelength, it travels via an essentially independent optical “carrier” wave within the waveguide, so that many signals can be transmitted simultaneously, or “multiplexed”, by the optical fiber. In the receiver, these signals can be separated by wavelength selective filters and routed to different detectors. Additional laser diodes may be used in the receiver as local oscillators (LO) for heterodyne detection of the optical signals. Let us now consider the advantages of an optical fiber interconnector like that shown in Fig. 1.1.

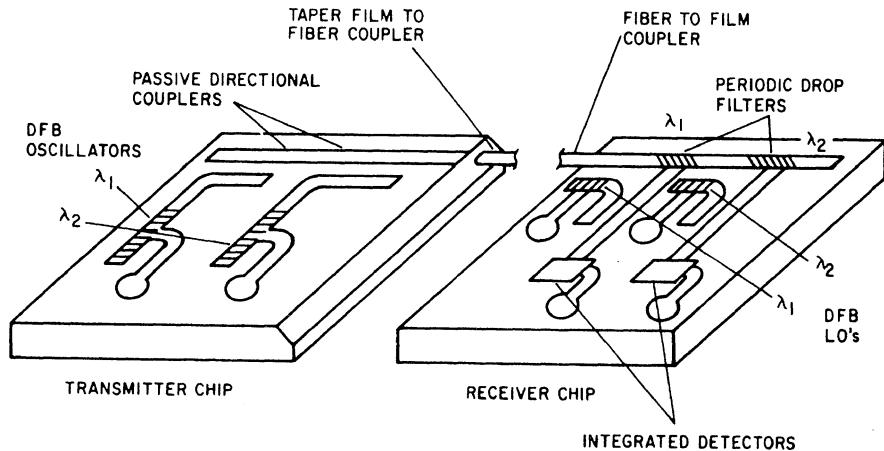


Fig. 1.1 Monolithic integrated optic system for optical communications

1.1.1 Comparison of Optical Fibers with Other Interconnectors

For many years, the standard means of interconnecting electrical subsystems, including integrated circuits, has been either the metallic wire or the radio link through the air. The optical fiber waveguide has many advantages over either of these conventional methods. The most important of these have been listed in Table 1.1 for easy reference, and they are discussed further below.

In modern electronic systems, such as those found in aircraft or in ground based communications systems, it is often necessary to run bundles of wires over considerable distances. These wires can act as receiving antennas, in which extraneous signals are generated by induction from the electromagnetic fields that surround the wire. These fields may be, for example, stray fields from adjacent wires, radio waves in the surrounding environment, or gamma radiation released during a nuclear explosion. In such applications as airborne radar, missile guidance, high-voltage power line fault sensing, and multichannel telecommunications, it is critically important that the system continue to operate normally in the presence

Table 1.1 Comparative evaluation of optical interconnectors

<i>Advantages</i>
Immunity from electromagnetic interference (EMI)
Freedom from electrical short circuits or ground loops
Safety in combustible environment
Security from monitoring
Low-loss transmission
Large bandwidth (i.e., multiplexing capability)
Small size, light weight
Inexpensive, composed of plentiful materials

<i>Major disadvantage</i>
Difficult to use for electrical power transmission

of severe electromagnetic interference (EMI). Metallic wires can, of course, be shielded, as in the case of coaxial cables, but the metallic shield adds weight, is costly, and produces parasitic capacitance that limits the frequency response or the bandwidth. The optical fiber waveguide has inherent immunity to most forms of EMI, since there is no metallic wire present in which current can be induced by stray electromagnetic coupling. In addition, it is easy to exclude undesired light waves by covering the fiber (or fiber bundle) with an opaque coating. “Cross talk”, or interference between the signals carried on adjacent optical fibers in a bundle, is also minimal because each wave-guiding core of the fibers is surrounded by a relatively thick cladding, through which the fields of the guided optical wave cannot penetrate. The immunity that guided optical waves have to EMI is sufficient reason, of itself, to prefer the use of optical fiber interconnectors instead of wires or cables in many applications. However, as we shall see, there are also quite a few other advantages to be gained by their use.

Unlike metal wires, optical fibers do not allow flow of electrical current, so that electrical short circuits cannot occur. Thus, optical fibers can be bound into a tight bundle and can be routed through metallic conduits without concern for electrical insulation. Optical fiber interconnections are particularly useful in high-voltage applications, such as transmitting telemetry data from, and control signals to, power transmission lines and switchgear. In this case, the insulating properties of the optical fiber itself eliminate the need for expensive isolation transformers. Since no spark is created when an optical fiber breaks, these fibers can also be used advantageously in combustible, or even explosive, environments.

In military and other high-security applications, optical fibers provide more immunity from *tapping* or monitoring than do wires or radio links. This is because there is no electromagnetic field extending outside of the optical fiber from which a signal can be tapped by electromagnetic induction in a pickup loop or antenna. In order to tap a signal from an optical fiber waveguide, one must somehow pierce the light-confining cladding of the fiber. This is very difficult to do without disrupting the light-wave transmission within the fiber to an extent which can be detected, thus easily permitting appropriate countermeasures.

Perhaps the most important advantage of fibers is that they can be used for low-loss transmission of optical signals over 100 kilometer-length paths with bandwidths greater than 40 Gigabit/s, without optical amplifiers. An overall transmission rate of 1800 Gb/s has been achieved by the use of dense wavelength division multiplexing (DWDM) and erbium-doped fiber amplifiers (EDFA) to transmit 180 10 Gb/s channels over 7000 km [7]. The loss can be reduced to less than 2 dB/km, even in relatively inexpensive, multimode, commercially available fibers. In single-mode fibers attenuation of less than 0.2dB/km is common. The losses in fibers are relatively independent of frequency, while those of competitive interconnectors increase rapidly with increasing frequency. For example, the data shown in Fig. 1.2 indicate losses in twisted-pair cable, which is commonly used as an interconnector in avionic systems, increase substantially at modulation frequencies above about 100 kHz. Coaxial cables are useful for transmission over relatively short paths, at frequencies up to about 100 MHz, even though losses are large, but above that frequency

Fig. 1.2 Comparison of attenuation in optical fiber with that in twisted-pair cable

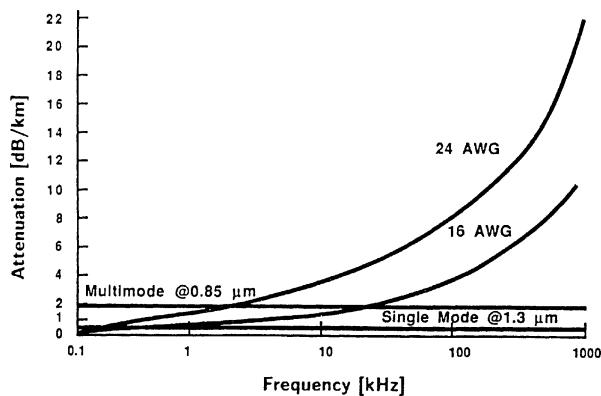
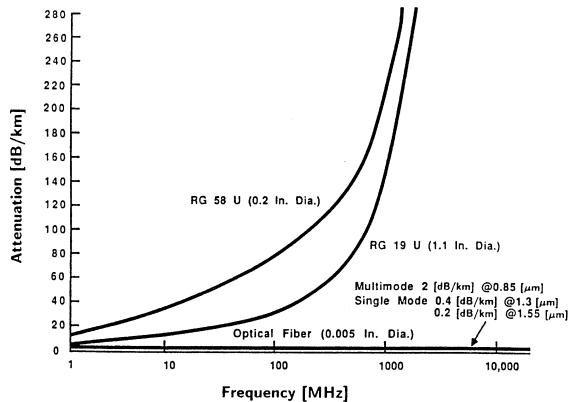


Fig. 1.3 Comparison of attenuation in optical fiber with that in coaxial cable



losses become excessive, as shown in Fig. 1.3. By comparison, attenuation in fibers is insignificant even at frequencies up to 10 GHz. The maximum frequency at which fibers can be used to transmit a signal is limited not by attenuation, per se, but rather indirectly by the phenomenon of dispersion.

The least expensive optical fibers that are readily available today are multimode fibers, in which the light waves propagate simultaneously in a number of different optical modes. Since each of these modes has a different group velocity, a pulse of light traveling along a multimode fiber is broadened. Transformation of this time-domain pulse broadening to the frequency domain results in a corresponding bandwidth product of about 1 GHz-km for currently available fibers. Modal dispersion can be avoided, of course, by using a single-mode fiber, in which the core diameter is made very small ($< 10 \mu\text{m}$ for visible or near-infrared wavelength) to cut off propagation of the higher-order modes. In that case, the bandwidth is limited only by the material dispersion, or variation of refractive index in the core with wavelength, and transmission of 1800 Gb/s over a 7000 km path has been achieved [7]. Single-mode

fibers are relatively expensive compared to the multimode type, and coupling and splicing problems are greatly aggravated by the small core diameter. An alternative approach is to use a graded index multimode fiber in which the refractive index of the core is graded from a maximum on axis to a minimum at the interface with the cladding. This grading of the index tends to equalize the effects of modal dispersion and it can be used to produce multimode fibers with bandwidths up to 3 GHz km [8, 9, 10].

In most applications, the large bandwidth of optical fibers will not be used to transmit a lone signal of that bandwidth, but rather to multiplex many signals of smaller bandwidth onto the same carrier light wave. This multiplexing capability, combined with the fact that the fiber diameter is typically hundreds of times smaller than that of a coaxial cable, means that the number of information channels per cross sectional area is on the order of 10^6 times larger when optical fibers are used. This is an important consideration when space is limited as in the case of aircraft, ships, or conduits under the streets of large cities, or undersea cables.

A number of ways in which optical fibers offer improved performance over their metallic counterparts have been described. However, the replacement of copper wires or metallic waveguides by fibers also can be an effective means of reducing cost in many applications. The basic materials from which optical fibers are fabricated, glasses and plastics, are more abundant and less costly than copper. Thus, once development costs are recovered, fibers are inherently less expensive to produce than copper wires or cables. In addition, fibers have a diameter hundreds of times smaller than that of a coaxial cable, thus resulting in less volume and weight. As mentioned previously, the wide bandwidth of fibers permits the multiplexing of thousands of information channels on each one, thus further reducing size and weight requirements. Savings in size and weight accomplished by using fibers instead of copper wires or cables can lead to a significant reduction of operating costs, as well as reducing the initial cost of system fabrication. For example, aircraft fuel consumption depends strongly on weight. A decrease of several hundred pounds in the total weight can mean thousands of dollars worth of fuel saved over the lifetime of the aircraft.

The one disadvantage of optical fibers compared to metallic wire, as noted in Table 1.1, is that it is difficult to transmit useful amounts of electrical power over them. However, it is not impossible to do so. Dentai, et al. [11] reported a long-wavelength optical-to-electrical generator that develops 10 volts output when illuminated with 1480–1650 nm lightwaves from a fiber.

In discussing the various advantages of optical fibers, we already noted that large signal bandwidths on the order of ten gigahertz could be transmitted. However, it is not sufficient to have the capability of transmitting such signals if we do not also possess the ability to generate and process them. Electrical integrated circuits offer little hope of operating at frequencies above about several gigahertz, because elements linked together by wires or other forms of metallic interconnections inevitably have stray inductances and capacitances associated with them that limit frequency response. Thus, we are led to the concept of an optical integrated circuit in which information is transported by a beam of light.

1.1.2 Comparison of Optical Integrated Circuits with Electrical Integrated Circuits

The optical integrated circuit has a number of advantages when compared to either its counterpart, the electrical integrated circuit, or to conventional optical signal processing systems composed of relatively large discrete elements. The major advantages of the OIC are enumerated in Table 1.2.

The optical integrated circuit inherently has the same large characteristic bandwidth as the optical fiber because, in both cases, the carrier medium is a lightwave rather than an electrical current. Thus, the frequency limiting effects of capacitance and inductance can be avoided. The design and fabrication of a large scale OIC with a bandwidth to match that of an optical fiber, while feasible in principle, probably will require many years of technology development. However, many practical applications of OIC's have already been accomplished (Chapter 20), and the future is promising.

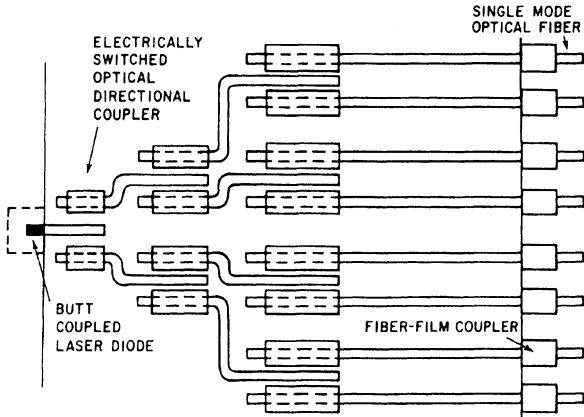
It should be possible to multiplex hundreds of signals onto one optical waveguide channel by using the frequency division multiplexing scheme that has been diagrammed in Fig. 1.1. A sixteen-channel DFB laser transmitter of this type has been produced by Rabon et al. [12], incorporating DFB lasers along with electroabsorption modulators on an InP substrate. The coupling of many optical signals into one waveguide can be conveniently and efficiently accomplished in the optical integrated circuit. Dual-channel directional couplers like those shown in Fig. 1.1 have been demonstrated to have coupling efficiency approaching 100% [13]. To implement the same type of bus access coupling to an optical fiber waveguide by using large discrete components such as the combiner-tap [14] would involve losses of at least 1 dB per insertion.

In addition to facilitating the coupling of many signals onto a waveguide, the optical integrated circuit also lends itself to convenient switching of signals from one waveguide to another. As shown diagrammatically in Fig. 1.4, this can be accomplished by using electro-optic switching. A deposited metal plate on top of, or between, the channels of a dual channel directional coupler can be used to

Table 1.2 Comparative evaluation of optical integrated circuits

<i>Advantages</i>
Increased bandwidth
Expanded frequency (wavelength) division multiplexing
Low-loss couplers, including bus access types
Expanded multipole switching (number of poles, switching speed)
Smaller size, weight, lower power consumption
Batch fabrication economy
Improved reliability
Improved optical alignment, immunity to vibration
<i>Major disadvantage</i>
High cost of developing new fabrication technology

Fig. 1.4 Optical integrated circuit directional coupler tree with electro-optic switching



control the transfer of optical power [15] Electro-optic switches of this type are now commercially available from a number of different suppliers.

When compared to larger, discrete component optical systems, OIC's can be expected to have the same advantages that electrical integrated circuits enjoy over hand-wired discrete component circuits. These include smaller size, weight and lower power requirements, as well as improved reliability and batch fabrication economy. In addition, optical alignment and vibration sensitivity, which are difficult problems in discrete component optical systems, are conveniently controlled in the OIC.

Enumeration of the many advantages of optical integrated circuits and of optical fiber interconnections suggests that integrated optics may completely replace conventional electronics, and indeed some persons do hold that view. However, the prevalent opinion is that it is more likely that integrated-optic systems will substantially augment, but not entirely replace, electronic ones. The relatively high cost of developing the technology of OIC's may probably limit their application to situations where the superior performance obtainable can justify the expense. Because optical fiber wave guides ultimately offer cost savings as well as performance improvements, they may very well replace wires and cables for most signal transmission. However, they cannot easily be used for low-frequency electrical energy transmission. Power transmission lines will most likely remain metallic. At present, integrated optics is still a rapidly developing field, and new applications are certain to emerge as the technology improves.

1.2 Substrate Materials for Optical Integrated Circuits

The choice of a substrate material on which to fabricate an optical integrated circuit depends most strongly on the function to be performed by the circuit. In most cases, the OIC may consist of a number of different optical devices such as sources,

modulators, and detectors, and no one substrate material will be optimum for all of them. Thus, a compromise must be made. The first step is to decided whether a hybrid or a monolithic approach is preferred, as elaborated below.

1.2.1 Hybrid Versus Monolithic Approach

There are two basic forms of optical integrated circuits. One of these is the hybrid, in which two or more substrate materials are somehow bonded together to optimize performance for different devices. The other is the monolithic OIC, in which a single substrate material is used for all devices. Since most OIC's will require a source of light, monolithic circuits can only be fabricated in optically active materials, such as the semiconductors listed in Table 1.3. Passive materials like quartz, lithium niobate or polymers are also useful as substrate materials but generally an external light source, such as a semiconductor laser, must somehow be optically and mechanically coupled to the substrate. However, in recent years significant progress has been made in producing light emitters and amplifiers by incorporating erbium and other atoms ions into passive substrate materials such as glasses and polymers, suggesting that it may be practical to make monolithic OIC's in them. OIC's of this type are discussed in Chapter 5.

While silicon is inherently a passive material, it can be modified by applying the techniques of nanophotonics to make it capable of light generation, and even laser action. Such silicon light emitters are described in Chapter 22.

The major advantage of the hybrid approach is that the OIC can be fabricated using existing technology, piecing together devices which have been substantially optimized in a given material. For example, one of the earliest OIC's to perform a complex system function was the rf spectrum analyzer, which combined a commercially available GaAAs diode laser and a silicon photodiode detector array with an acousto-optic modulator on a lithium niobate substrate [16, 17, 18]. A hybrid butt-coupling approach was used to efficiently couple both the laser diode and the detector array to the LiNbO₃ substrate. In this case, the hybrid approach made possible the combining of already well-developed technologies for GaAAs heterojunction lasers [19] LiNbO₃ acousto-optic waveguide modulators cite [20] and Si photodiode arrays [21].

Table 1.3 Substrate materials for optical integrated circuits

<i>Passive</i> (Incapable of light generation)	<i>Active</i> (Capable of light generation)
Quartz	gallium arsenide
lithium niobate	gallium aluminum arsenide
lithium tantalate	gallium arsenide phosphide
tantalum pentoxide	gallium indium arsenide
niobium pentoxide	other III-V & II-VI semiconductors
silicon	
polymers	

While the hybrid approach to OIC fabrication provides a convenient way for one to implement many desired functions, it has the disadvantage that the bonds holding the various elements of the circuit together are subject to misalignment, or even failure, because of vibration and thermal expansion. Also, the monolithic approach is ultimately cheaper if mass production of the circuit is desired, because automated batch processing can be used. This yields a low perunit cost after design and development costs are recovered. For these reasons, monolithic OIC's are likely to become the most common type in use once the technology has matured, even though the first commercially available OIC's have been produced using the hybrid approach.

1.2.2 III-V and II-VI Ternary Systems

Most monolithic OIC's can be fabricated only in active substrates, in which light emitters can be formed. This essentially limits the choice of materials to semiconductors, such as those listed in Table 1.3. The III-V (or II-VI) ternary or quaternary compounds are particularly useful because the energy bandgap of the material can be changed over a wide range by altering the relative concentrations of elements. This feature is very important to the solution of one of the basic problems of monolithic OIC fabrication. Semiconductors characteristically emit light at a wavelength corresponding approximately to their bandgap energy. They also very strongly absorb light having a wavelength less than, or equal to, their bandgap wavelength. Thus, if a light emitter, waveguide, and detector are all fabricated in a single semiconductor substrate such as GaAs, light from the emitter will be excessively absorbed in the waveguide but not absorbed strongly enough in the detector. As will be explained in detail in following chapters, the composition of a ternary or quaternary material can be adjusted in the various component devices of the OIC to effectively eliminate these effects.

So far, most of the research in monolithic OIC's has used the gallium aluminum arsenide, $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$, system or the gallium indium arsenide phosphide, $\text{Ga}_x\text{In}_{(1-x)}\text{As}_{(1-y)}\text{P}_y$, system. These materials, often designated by the abbreviated formulae GaAlAs and GaInAsP, have a number of properties that make them especially useful for OIC fabrication. The most important of these have been enumerated in Table 1.4. By changing the fractional atomic concentration of the constituents, the emitted wavelength can be varied from $0.65\mu\text{m}$ (for AlAs) to $1.7\mu\text{m}$ (for GaInAsP). GaAlAs and GaInAsP also have relatively large electro-optic and acousto-optic figures of merit, making them useful for optical switch and modulator fabrication. Widespread use of GaAlAs and GaInAsP has greatly reduced their cost as compared to that of other III-V or II-VI compounds. GaAs wafers suitable for use as OIC substrates can be purchased from a number of commercial suppliers.

The fact that GaAlAs technology is already well developed compared to that of other III-V ternaries make its use advantageous. It also has the unique property that the lattice constants of GaAs and AlAs are almost identical (5.646 and 5.66\AA , respectively [22]. Thus, layers of GaAlAs with greatly different Al concentrations can be epitaxially grown on top of each other with minimal interfacial lattice strain

Table 1.4 Properties of GaAs, GaAlAs and GaInAsP useful in optical integrated circuits

Transparency	0.6–12 μm
Emitted wavelength	0.65–1.7 μm
Lattice matching	Negligible lattice mismatch results in minimal strain
Switching	Large electro-optic and acusto-optic figures of merit $n_0^3 r_{41} \simeq 6 \times 10^{-11} \text{ m/V}$ $M = \frac{n_0^6 p^2}{\varrho v_s^3} \simeq 10^{-13} \text{ s}^3/\text{kg}$
Technology	Epitaxy, doping, ohmic, contacts, masking, etching all are well developed
Cost	Less than other III–V or II–VI materials

being introduced. This is particularly important in the fabrication of multilayered, heterojunction lasers, as will be discussed in Chapter 14. No other III–V or II–VI pair has lattice constants that are matched as well as those of GaAs and AlAs. As a result, interfacial strain is a major problem in the fabrication of multilayered devices in these materials. However, a strong desire to produce light sources emitting at wavelengths longer than those achievable with GaAlAs has led to the development of GaInAsP lattice matching technology capable of yielding highly efficient laser diodes emitting at wavelengths of 1.3 μm and 1.55 μm . These wavelengths correspond to minimum values of absorption and dispersion in glass optical fibers used for long distance telecommunications.

Work on GaAlAs and GaInAsP monolithic OIC's has already produced many of the desired optical devices in the required monolithic form. However, in most cases the level of integration so far has been limited to relatively few devices per chip. Many of the monolithically integrated devices have performed as well as, or better than, their discrete counterparts. The lower levels of integration that have been achieved so far appear to stem from the personal preferences of the researchers involved rather than from any fundamental technical limitation. Some persons have preferred to work on heterojunction lasers, while others have concentrated on waveguides, or detectors. This specialization has been beneficial in that it has led to the invention and development of relatively sophisticated embodiments of many of the devices required in OIC's. Such devices as heterojunction, distributed feedback lasers, acousto-optic waveguide modulators, and electro-optically switched directional couplers are examples of well developed technology. In the recent years, more researchers have turned to the problem of actually integrating a number of devices on one substrate. This has led to monolithic OIC's, combining lasers, modulators or switches, and detectors [23–31]. Such a trend no doubt will continue in the future and should result in larger scale integration.

1.2.3 Hybrid OIC's in LiNbO_3

While much has been accomplished in the realm of monolithic OIC's, they have been slow to achieve commercial availability because of the very complex technology required for their fabrication. By contrast, hybrid OIC's fabricated in LiNbO_3

Table 1.5 Properties of LiNbO₃ useful in optical integrated circuits

Transparency	0.2–12 μm low loss
Emitted wavelength	none
Switching	Large electro-optic and acusto-optic figures of merit $n_0^3 r_{33} \simeq 3 \times 10^{-10} \text{ m/V}$ $M = \frac{n_0^6 b^2}{Q v_i^3} \simeq 7 \times 10^{-15} \text{ s}^3/\text{kg}$
Technology	Waveguide fabrication, masking, etching, polishing all are well developed
Cost	More than GaAs

are presently commercially available from a number of different suppliers. These are relatively simple, including electro-optic modulators, Mach-Zehnder interferometric modulators, electro-optic switches and optical amplifiers. These OIC's take advantage of the wide wavelength range of transparency and large electro-optic coefficient of LiNbO₃, as well as some other beneficial properties of that material which are listed in Table 1.5.

1.2.4 Organization of this Book

This book has been planned so as to serve both as a text for a course in integrated optics, and as a readable reference volume for scientists and engineers in industry. Hence, Chapter 1 has given an overview of the goals and basic principles of integrated optics and has compared this new approach to those alternative techniques now in use. The close relationship and interdependence between optical fiber waveguides and optical integrated circuits has been described. Thus, we are prepared to begin a detailed study of the component elements of an optical integrated circuit.

Chapters 2 and 4 are devoted to the fundamental element that literally ties the OIC together: the optical waveguide. Without effective, low-loss optical waveguides one cannot even contemplate an OIC. Hence, it is logical to begin the study of integrated optics by considering the waveguide. The theory of optical waveguides is discussed in Chapters 2 and 3, beginning with the basic three-layer planar waveguide structure. A comparison is made between the geometric or “ray optic” approach and the electromagnetic field or “physical optic” approach. Theoretical derivations of the optical mode profiles and cutoff conditions for common waveguide geometry's are compared with corresponding experimental results. Following the development of the theory of optical waveguides, techniques of waveguide fabrication are discussed in Chapter 4.

Chapter 5 covers the emerging area of polymer and fiber integrated optics. As the commercial market for integrated optic devices has grown, the relatively low cost of polymers and glasses has made them particularly attractive for use in OIC's.

In Chapter 6, optical losses in waveguides are described, and techniques used to measure those losses are discussed.

Once the design and fabrication of the optical waveguide has been considered, the next immediate question is that of efficiently coupling optical energy into or out of a waveguide. Chapter 7 describes a number of ways for accomplishing this aim. Since much of the laboratory research in OIC's involves the use of conventional (non-integrated) lasers, we consider prism and grating couplers, which are suitable for the beams of light emitted by such sources. Additionally, hybrid couplers like the transverse (end-butt) coupler and various fiber-to-waveguide couplers are described. Finally, in Chapter 8 fully integrated, monolithic waveguide-to-waveguide couplers like the dual-channel directional coupler are dealt with. This latter type of coupler is composed of two waveguiding channels spaced closely enough so that the evanescent "tails" of the optical modes overlap, thus producing coherent coupling of optical energy, in a manner analogous to that of the familiar slotted waveguide used as a coupler at microwave frequencies. As discussed in Chapter 9, the dual-channel directional coupler can also be used as an optical modulator or switch, because the strength of coupling between channels can be electrically controlled by adding properly designed electrodes.

The topic of optical beam modulation is very important to integrated optics because the wide bandwidths of optical fibers and OIC's can be utilized only if wide-band modulators are available to impress the signal information onto the optical beam. Hence, Chapters 9 and 10 are devoted to detailed considerations of the various types of optical modulators that can be fabricated in a monolithic embodiment. These can be divided generally into two classes, electro-optic and acousto-optic, depending on whether the electrical signal is coupled to the optical beam by means of electro-optic change of the index of refraction, or by changes in index produced by acoustic waves. Of course, in the latter case, the electrical signal must first be coupled to the acoustic wave by an acoustic transducer.

While discrete (non-integrated) laser sources are adequate for laboratory research, and for some integrated optics applications, most of the benefits of an OIC technology cannot be obtained without using a monolithically integrated light source. The prime choice for such a source is the p-n junction laser diode, fabricated on a semiconductor substrate that also supports the rest of the optical integrated circuit. Because of the critical importance of this device to the field of integrated optics, and because of its complexity, Chapters 11, 12, 13, 14 and 15 have been dedicated to a thorough, systematic development of semiconductor laser diode theory and technology. Chapter 11 reviews the basic principles of light emission in semiconductors, explaining the phenomena of spontaneous and stimulated emission of photons. The important effect of the energy band structure on the emission process is discussed. In Chapter 12, attention is focused on specific types of semiconductor laser structures. The concept of optical field confinement is introduced and the development of the confined-field laser is traced up to the modern heterostructure diode laser, which has become the standard light source in OIC's. The phenomenon of spontaneous emission is important not only in lasers but also in optical amplifiers. Chapter 13 provides coverage of the topic of optical amplifiers, which have emerged as very important components of optical communications systems. Chapter 14 provides a detailed description of heterojunction, confined-field lasers. Finally, in Chapter 15,

the distributed feedback(DFB) laser is discussed. This device features a diffraction grating that provides the optical feedback necessary for laser oscillation, thus eliminating the need for cleaved end-faces or other reflecting mirrors to form an optical cavity. Since efficient reflecting end surfaces are difficult to fabricate in a planar monolithic OIC, the DFB laser is an attractive alternative.

Chapter 16 may appear at first to be out of place, since it deals with the direct modulation of semiconductor lasers. However, the technique of directly modulating the light output of the laser diode by varying the input current is distinct from either electro-optic or acousto-optic modulation methods described in Chapters 9 and 10 in that it cannot be understood without a knowledge of the characteristics of light generation in the semiconductor laser. Hence, presentation of the material in Chapter 16 has been postponed until after the explanation of the laser diode. Input current modulation rates that are desired in OIC's. Hence, a thorough discussion has been given of the various techniques that have been used and the results that have been obtained to date.

Although the signal processing in an OIC is done in the optical domain, the output information is most often desired in the form of an electrical signal to interface with electronic systems. This requires an optical to electrical transducer, commonly called a photodetector. In Chapter 17 we describe various types of photodetectors that can be fabricated as elements of a monolithic OIC. The advantage of a waveguide detector as compared to the corresponding bulk device are enumerated. Also, techniques of locally modifying the effective bandgap of the substrate in the vicinity of the detector to increase conversion efficiency are described.

A recent development in the field of integrated optics has been the development of “quantum well” devices, in which a multiplayer structure of very thin layers is used to create a “Superlattice”. This superlattice structure yields devices (e.g. lasers, detectors, modulators and switches) with improved operating characteristics, Chapter 18 is devoted to a review of these devices.

To complete the study of integrated optics presented in this book, a number of applications of integrated optic devices and systems are described in Chapters 19, 20, 21 and 22, and trends in the field are analyzed. Applications in the new area of micro-electro-optical-mechanical devices (MEM's) are discussed, as well as those in the long-established telecommunications area. Chapters 19 and 21 review applications of optoelectronic devices in the new fields of MEM's and Wireless systems.

A new Chapter 22 has been added to this edition to explore the latest area of integrated optics, which is called “nanophotonics.” The advantages of the newly developed ability to form nanometer-sized features on semiconductor chips and other substrates go far beyond just reducing the size of the OIC or PIC. When feature dimensions approach the same size as the wavelength of the light in the substrate material it can behave as a new “photonic crystal” with significantly different optical properties than those of the original material. Unique waveguides, couplers, switches, and light emitters can all be formed in photonic crystals.

Since integrated optics is a young field compared to that of electrical integrated circuits, it is difficult to envision exactly what course will be followed. However, it

seems certain that many significant contributions to the technology of optical signal generation, transmission and processing will result.

Problems

- 1.1 Name four advantages of optical fiber interconnects as compared to metallic wires.
- 1.2 Name four advantages of optical integrated circuits as compared to electrical integrated circuits.
- 1.3 What is the difference between a hybrid OIC and a monolithic OIC? What are the advantages and disadvantages of each type?
- 1.4 Why are GaAs, GaAlAs and GaInAsP particularly useful materials for the fabrication of optical integrated circuits?

References

1. T. Tamir (ed.): *Integrated Optics*, 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979)
2. S.E. Miller, A.G. Chynoweth: *Optical Fiber Communications* (Academic, New York 1979) Chap. 1
3. OSA Topical Meeting on Integrated Optics – Guided Waves, Materials and Devices, Las Vegas, NV (1972)
4. OSA Topical Meeting on Integrated Optics – Guided Waves, Materials and Devices, Las Vegas, NV (1974)
5. R.D. Maurer: Properties of research fibers for optical communications. OSA Topical Meeeting on Integrated Optics, New Orleans, LA (1974)
6. S.E. Miller: Optical-fiber transmission research. OSA Topical Meeting on Integrated Optics, New Orleans, LA (1974)
7. J.A. Cai, M. Nissov, A.N. Pilipetskii, A.J. Lucero, C.R. Davidson, D. Foursa, H. Kidorf, M.A. Mills, R. Menges, P.C. Corbett, D. Sutton, N.S. Bergano: 2.4 Tb/s (120×20 Gb/s) transmission over transoceanic distance using optimum fec overhead and 48% spectral efficiency. Optical Fiber Communication Conference and Exhibit, 2001. OFC 2001, Anaheim, CA (2001)
8. T.K. Woodward, S. Hunsche, A.J. Ritger, J.B. Stark: IEEE Phot. Techn. Lett. **11**, 382 (1999)
9. L.G. Cohen, P. Kaiser, C. Lin: IEEE Proc. **68**, 1203 (1980)
10. G. Giaretta, W. White, M. Wegmuller, T. Onishi: IEEE Phot. Techn. Lett. **12**, 347 (2000)
11. A.G. Dentai, C.R. Giles, E. Burrows, C.A. Burus, L. Stulz, J. Centanni, J. Hoffman, B. Moyer: IEEE Phot. Techn. Lett. **11**, 114 (1999)
12. G. Rabon: Proc. SPIE **2684**, 102 (1996)
13. S. Somekh, E. Garmire, A. Yariv, H. Garyin, R.G. Hunsperger: Appl. Opt. **13**, 327 (1974)
14. J.F. Dalgleish: IEEE Proc. **68**, 1226 (1980)
15. M. Papuchon: Appl. Phys. Lett. **27**, 289 (1975)
16. M.C. Hamilton, D.A. Wille, M.J. Miceli: Opt. Eng. **16**, 475 (1977)
17. D. Mergerian, E.C. Malarkey: Microwave J. **23**, 27 (September 1980)
18. D. Mergerian, E.C. Malarkey, R.P. Pautieu, S.C. Bradly, M. Mill, C.W. Baugh, A.L. Kellner, M. Mentzer: SPIE Proc. **321**, 149 (1982)
19. H. Kressel, M. Ettenberg: J. Appl. Phys. **47**, 3533 (1976)

20. C.S. Tsai: IEEE Trans. CAS-**26**, 1072 (1980)
21. C.L. Chen, J.T. Boyd: Channel waveguide array coupled to an integrated charge-coupled device (CCD). OSA Topical Meeting on Integrated Optics, Salt Lake City, UT (1978)
22. G. Giesecke: Lattice constants. *Semiconductors and Semimetals* 2, 63–75 (Academic, New York 1976), in particular, ps. 68 and 69
23. K. Sato, S. Sekine, Y. Kondo, M. Yamamoto: IEEE J. QE-**29**, 180 (1993)
24. P. Hvertas, G. Mier, M. Dotor, J. Anguita, D. Golmayo, F. Briones: Sensors and Actuators A: Physical **37**, 512 (1993)
25. S. Ura, M. Shinohara, T. Suhara, N. Nishihara: IEEE Phot. Techn. Lett. **6**, 239 (1994)
26. S. Lee, I. Jong, C. Waqng, C. Pien, T. Shih: IEEE J. Selected Topics in Quantum Electronics **6**, 197 (2000)
27. J. Ahadian, C. Fonstad, Jr.: Optical Engin. **37**, 3167 (1998)
28. H.-G. Bach, A. Umbach, S. van Waasen, R.M. Bertenberg, G. Unterborsch: IEEE J. Selected Topics in Quantum Electronics **2**, 418 (1996)
29. J.E. Johnson, J.-P. Ketelsen, J.A. Grenko, S.K. Sputz, J. Vandenberg, M.W. Focht, D.V. Stampaone, L.J. Petricolas, L.E. Smith, K.G. Glogovsky, G.J. Przybylek, S.N.G. Chu, J.L. Lentz, N.N. Tzafaras, L.C. Luther, T.L. Pernell, F.S. Walters, D.M. Romero, J.M. Freund, C.L. Reynolds, L.A. Gruezke, R. People, M.A. Alam: IEEE J. Selected Topics in Quantum Electronics **6**, 19 (2000)
30. L.H. Spiekman, J.M. Wiesenfeld, U. Koren, B.I. Miller, M.D. Chien: IEEE Photonic. Tech. Lett. **10**, 1115 (1998)
31. A. Krishnamoorthy, K. Goosen: IEEE J. Selected Topics in Quantum Electronics **4**, 899 (1998)

Chapter 2

Optical Waveguide Modes

The optical waveguide is the fundamental element that interconnects the various devices of an optical integrated circuit, just as a metallic strip does in an electrical integrated circuit. However, unlike electrical current that flows through a metal strip according to Ohm's law, optical waves travel in the waveguide in distinct optical modes. A mode, in this sense, is a spatial distribution of optical energy in one or more dimensions that remains constant in time. In this chapter, the concept of optical modes in a waveguiding structure is discussed qualitatively, and key results of waveguide theory are presented with minimal proof to give the reader a general understanding of the nature of light propagation in an optical waveguide. Then, in Chap. 3, a mathematically sound development of waveguide theory is given.

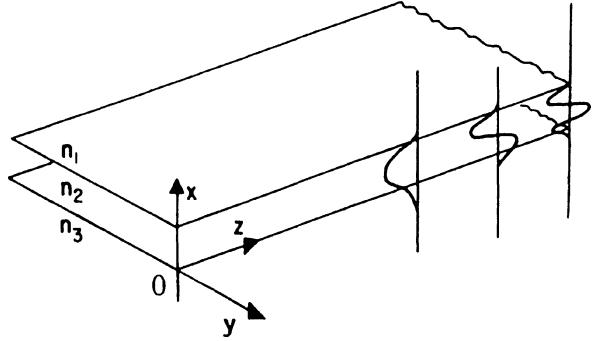
2.1 Modes in a Planar Waveguide Structure

As shown in Fig. 2.1, a planar waveguide is characterized by parallel planar boundaries with respect to one (x) direction, but is infinite in extent in the lateral directions (z and y). Of course, because it is infinite in two dimensions, it cannot be a practical waveguide for optical integrated circuits, but it forms the basis for the analysis of practical waveguides of rectangular cross section. It has therefore been treated by a number of authors, including McWhorter [1], McKenna [2], Tien [3], Marcuse [4], Taylor and Yariv [5] and Kogelnik [6]. In Section 2.1.2 we follow the approach of Taylor and Yariv [5] to examine the possible modes in a planar waveguide, without fully solving the wave equation.

2.1.1 Theoretical Description of the Modes of a Three-Layer Planar Waveguide

To begin the discussion of optical modes, consider the simple three-layer planar waveguiding structure of Fig. 2.1. The layers are all assumed to be infinite in extent in the y and z directions, and layers 1 and 3 are also assumed to be semi-infinite in the x direction. Light waves are assumed to be propagating in the z direction. It has

Fig. 2.1 Diagram of the basic three-layer planar waveguide structure. Three modes are shown, representing distributions of electric field in the x direction



been stated previously that a mode is a spatial distribution of optical energy in one or more dimensions. An equivalent mathematical definition of a mode is that it is an electromagnetic field which is a solution of Maxwell's wave equation

$$\nabla^2 \mathbf{E}(\mathbf{r}, t) = [n^2(\mathbf{r})/c^2] \partial^2 \mathbf{E}(\mathbf{r}, t)/\partial t^2, \quad (2.1)$$

where \mathbf{E} is the electric field vector, \mathbf{r} is the radius vector, $n(\mathbf{r})$ is the index of refraction, and c is the speed of light in a vacuum. For monochromatic waves, the solutions of (2.1) have the form

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}) e^{i\omega t}, \quad (2.2)$$

where ω is the radian frequency. Substituting (2.2) into (2.1) we obtain

$$\nabla^2 \mathbf{E}(\mathbf{r}) + k^2 n^2(\mathbf{r}) \mathbf{E}(\mathbf{r}) = 0, \quad (2.3)$$

where $\kappa \equiv \omega/c$. If we assume, for convenience, a uniform plane wave propagating in the z direction, i.e., $\mathbf{E}(\mathbf{r}) = E(x, y) \exp(-i\beta z)$, β being a propagation constant, then (2.3) becomes

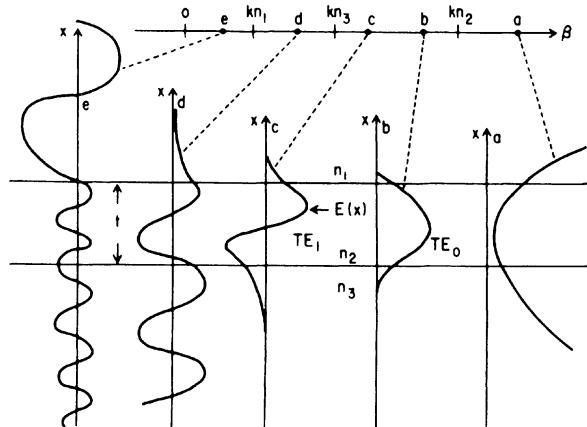
$$\partial^2 E(x, y)/\partial x^2 + \partial^2 E(x, y)/\partial y^2 + [k^2 n^2(\mathbf{r}) - \beta^2] E(x, y) = 0. \quad (2.4)$$

Since the waveguide is assumed infinite in the y direction, by writing (2.4) separately for the three regions in x , we get

$$\begin{aligned} \text{Region 1} \quad & \partial^2 E(x, y)/\partial x^2 + (k^2 n_1^2 - \beta^2) E(x, y) = 0 \\ \text{Region 2} \quad & \partial^2 E(x, y)/\partial x^2 + (k^2 n_2^2 - \beta^2) E(x, y) = 0 \\ \text{Region 3} \quad & \partial^2 E(x, y)/\partial x^2 + (k^2 n_3^2 - \beta^2) E(x, y) = 0, \end{aligned} \quad (2.5)$$

where $E(x, y)$ is one of the Cartesian components of $\mathbf{E}(x, y)$. The solutions of (2.5) are either sinusoidal or exponential functions of x in each of the regions, depending

Fig. 2.2 Diagram of the possible modes in a planar waveguide [2.5]



on whether $(k^2 n_i^2 - \beta^2)$, $i = 1, 2, 3$, is greater than or less than zero. Of course, $E(x, y)$ and $\partial E(x, y)/\partial x$ must be continuous at the interface between layers. Hence the possible modes are limited to those shown in Fig. 2.2.

Consider how the mode shape changes as a function of β , for the case of constant frequency ω and $n_2 > n_3 > n_1$. This relative ordering of the indices is quite a common case, corresponding, for example, to a waveguiding layer of index n_2 formed on a substrate with smaller index n_3 , surrounded by air of index n_1 . As we will see in Chapter 3, it is a necessary condition for waveguiding in Layer 2 that n_2 be greater than both n_1 and n_3 . When $\beta > kn_2$, the function $E(x)$ must be exponential in all three regions and only the mode shape shown as (a) in Fig. 2.2 could satisfy the boundary conditions of $E(x)$ and $\partial E(x)/\partial x$ being continuous at the interfaces. This mode is not physically realizable because the field increases unboundedly in Layers 1 and 3, implying infinite energy. Modes (b) and (c) are well confined guided modes, generally referred to as the zeroth order and first order transverse electric modes, TE_0 and TE_1 [7]. For values of β between kn_2 and kn_3 such modes can be supported. If β is greater than kn_1 but less than kn_3 , a mode like that in (d) will result. This type of mode, which is confined at the air interface but sinusoidally varying at the substrate, is often called a *substrate radiation mode*. It can be supported by the waveguide structure, but because it is continually losing energy from the waveguiding Region 2 to the substrate Region 3 as it propagates, it tends to be damped out over at short distance. Hence it is not very useful in signal transmission, but, infact, it may be very useful in coupler applications such as the tapered coupler. This type of coupler will be discussed in Chapter 6. If β is less than kn_1 the solution for $E(x)$ is oscillatory in all three regions of the waveguide structure. These modes are not guided modes because the energy is free to spread out of the waveguiding Region 2. They are generally referred to as the air radiation modes of the waveguide structure. Of course, radiation is also occurring at the substrate interface.

2.1.2 Cutoff Conditions

We shall see in Chapter 3, when (2.1) is formally solved, subject to appropriate boundary conditions at the interface, that β can have any value when it is less than kn_3 , but only discrete values of β are allowed in the range between kn_3 and kn_2 . These discrete values of β correspond to the various modes $TE_j, j = 0, 1, 2, \dots$ (or $TM_k, k = 0, 1, 2, \dots$). The number of modes that can be supported depends on the thickness t of the waveguiding layer and on ω, n_1, n_2 and n_3 . For given t, n_1, n_2 , and n_3 there is a cutoff frequency ω_c below which waveguiding cannot occur. This ω_c corresponds to a long wavelength cutoff λ_c .

Since wavelength is often a fixed parameter in a given application, the cutoff problem is frequently stated by asking the question, “for a given wavelength, what indices of refraction must be chosen in the three layers to permit waveguiding of a given mode?” For the special case of the so-called asymmetric waveguide, in which n_1 is very much less than n_3 , it can be shown (Chapter 3) that the required indices of refraction are related by

$$\Delta n = n_2 - n_3 = (2m + 1)^2 \lambda_0^2 / (32n_2 t^2), \quad (2.6)$$

where the mode number $m = 0, 1, 2, \dots$, and λ_0 is the vacuum wavelength. The change in index of refraction required for waveguiding of the lower-order modes is surprisingly small. For example, in a gallium arsenide waveguide with n_2 equal to 3.6 [8] and with t on the order of λ_0 , (2.6) predicts that a Δn on the order of only 10^{-2} is sufficient to support waveguiding of the TE_0 mode.

Because only a small change in index is needed, a great many different methods of waveguide fabrication have proven effective in a variety of substrate materials. The more important of these have been listed in Table 2.1 so that the reader will be familiar with the names of the techniques when they are mentioned in the following discussion of experimental observations of waveguide performance. A thorough explanation of the methods of waveguide fabrication is given in Chapters 4 and 5.

Table 2.1 Methods of fabricating waveguides for optical integrated circuits

-
- 1) Deposited thin films (glass, nitrides, oxides, organic polymers)
 - 2) Photoresist films
 - 3) Ion bombarded glass
 - 4) Diffused dopant atoms
 - 5) Heteroepitaxial layer growth
 - 6) Electro-optic effect
 - 7) Metal film stripline
 - 8) Ion migration
 - 9) Reduced carrier concentration in a semiconductor
 - a) epitaxial layer growth
 - b) diffusion counterdoping
 - c) ion implantation counterdoping or compensation
-

2.1.3 Experimental Observation of Waveguide Modes

Since the waveguides in optical integrated circuits are typically only a few micrometers thick, observation of the optical mode profile across a given dimension cannot be accomplished without a relatively elaborate experimental set-up, featuring at least $1000\times$ magnification. One such system [9], which works particularly well for semiconductor waveguides, is shown in Fig. 2.3. The sample, with its waveguide at the top surface, is fixed atop an x - y - z micropositioner. Microscope objective lenses, used for input beam coupling and output image magnification, are also mounted on micropositioners to facilitate the critical alignment that is required. The light source is a gas laser, emitting a wavelength to which the waveguide is transparent. For example, a helium-neon laser operating at $1.15\text{ }\mu\text{m}$ is good for GaAs, GaAlAs and GaP waveguides, while one emitting at 6328 \AA can be used for GaP but not for GaAlAs or GaAs. For visual observation of the waveguide mode, the output face of the waveguide can be imaged onto either a white screen or an image converter (IC) screen depending on whether visible wavelength or infrared (ir) light is used. The lowest order mode ($m = 0$) appears as a single band of light, while higher order modes have a correspondingly increased number of bands, as shown in Fig. 2.4.

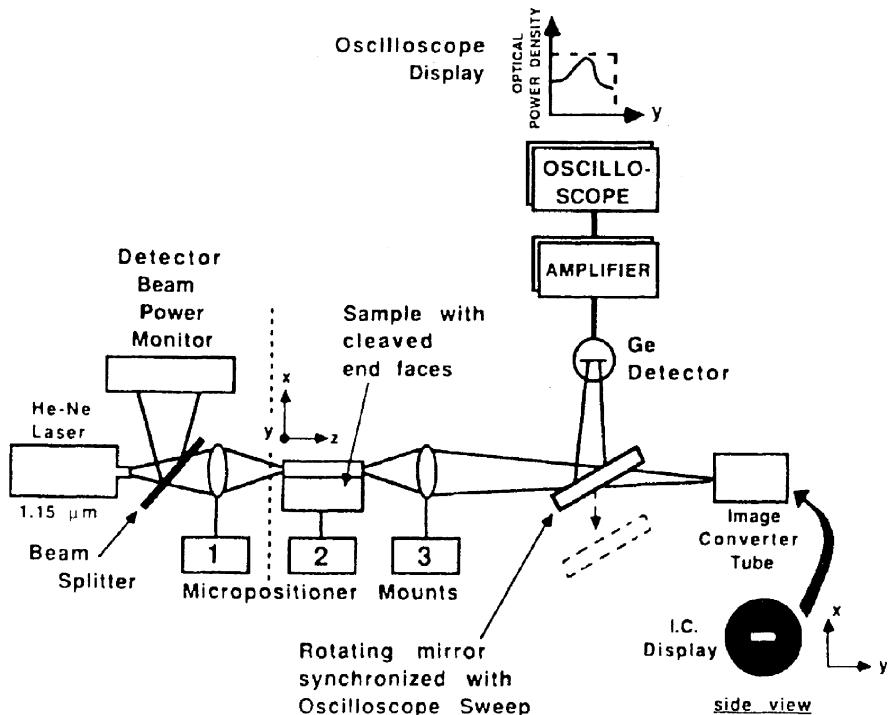


Fig. 2.3 Diagram of an experimental setup than can be used to measure optical mode shapes [2.9]

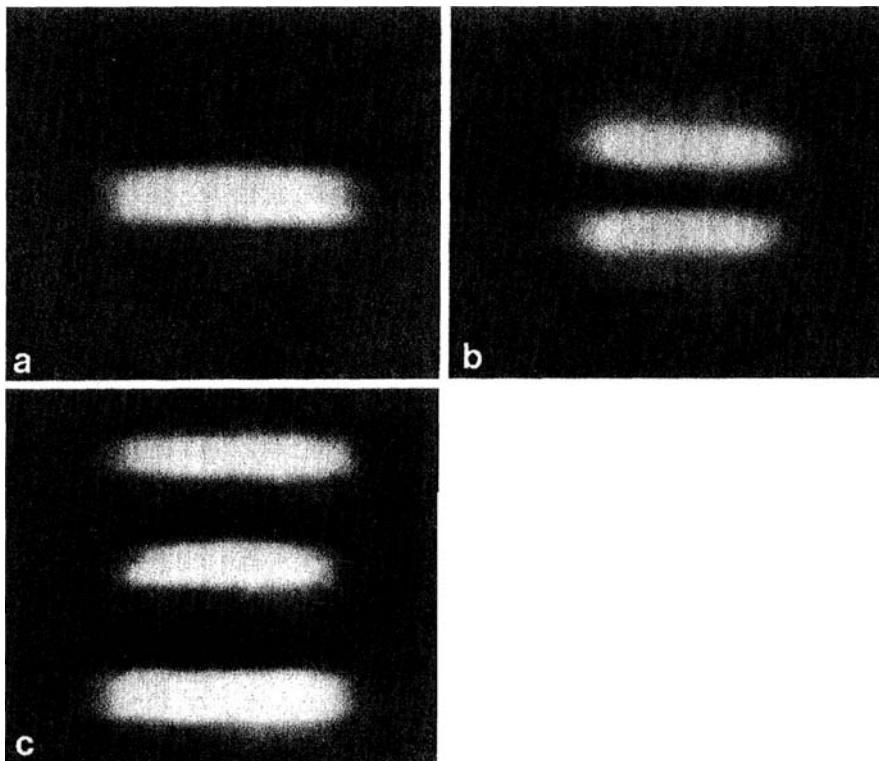
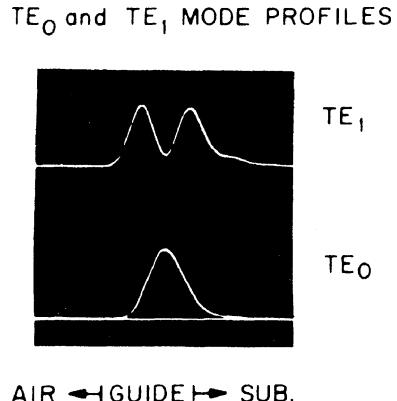


Fig. 2.4a, b, c Optical mode patterns in a planar waveguide, **a** TE_0 , **b** TE_1 , **c** TE_2 . In the planar guide, light is unconfined in the y direction, and is limited, as shown in the photos, only by the extent of spreading of the input laser beam. For the corresponding TE_{xy} patterns of a rectangular waveguide, see [10]

The light image appears as a band rather than a spot because it is confined by the waveguide only in the x direction. Since the waveguide is much wider than it is thick the laser beam is essentially free to diverge in the y direction.

To obtain a quantitative display of the mode profile, i.e. optical power density vs. distance across the face of the waveguide, a rotating mirror is used to scan the image of the waveguide face across a photodetector that is masked to a narrow slit input. The electrical signal from the detector is then fed to the vertical scale of an oscilloscope on which the horizontal sweep has been synchronized with the mirror scan rate. The result is in the form of graphic displays of the mode shape, like those shown in Fig. 2.5. Note that the modes have the theoretically predicted sinusoidal-exponential shape, by remembering that what is observed is optical power density, or intensity, which is proportional to E^2 . Details of the mode shape, like the rate of exponential decay (or extinction) of the evanescent “tail” extending across the waveguide-substrate and waveguide-air interfaces, depend strongly on the values of Δ at the interface. As can be seen in Fig. 2.5, the extinction is much sharper at

Fig. 2.5 Optical mode shapes are measured using the apparatus of Fig. 2.3. The waveguide in this case was formed by proton implantation into a gallium arsenide substrate to produce a $5\ \mu\text{m}$ thick carrier-compensated layer [12]



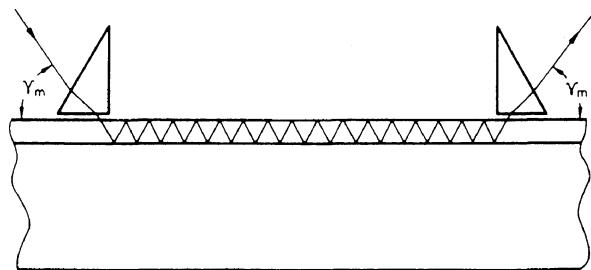
the waveguide-air interface where $\Delta n \simeq 2.5$ than at the waveguide-substrate plane where $\Delta n \simeq 0.01\text{--}0.1$.

A system like that shown in Fig. 2.3 is particularly useful for analysis of mode shapes in semiconductor waveguides, which generally support only one or two modes because of the relatively small Δn at the waveguide-substrate interface. Generally, the position of the focused input laser beam can be moved toward the center of the waveguide to selectively pump the zeroth order mode, or toward either the air or substrate interface to select the first order mode. It becomes very difficult to visually resolve the light bands in the case of higher-order, multimode waveguides because of spatial overlapping, even though the modes may be electromagnetically distinct and non-coupled one to another. Waveguides produced by depositing thin films of oxides, nitrides or glasses onto glass or semiconductor substrates usually are multi-mode, supporting 3 or more modes, because of the larger waveguide-substrate Δn [11–14]. For waveguides of this type, a different experimental technique, employing prism coupling, is most often used to analyze the modes.

The prism coupler will be discussed in detail in Chapter 7. At this point it suffices to say that the prism coupler has the property that it selectively couples light into (or out of) a particular mode, depending on the angle of incidence (or emergence). The mode-selective property of the prism coupler, which is illustrated in Fig. 2.6, results from the fact that light in each mode within a waveguide propagates at a different velocity, and continuous phase-matching is required for coupling. The particular angle of incidence required to couple light into a given mode or the angle of emergence of light coupled out of a given mode can both be accurately calculated from theory, as will be seen in Chapter 7. The prism coupler can thus be used to analyze the modes of a waveguide. This can be done in two ways.

In one approach, the angle of incidence of a collimated, monochromatic laser beam on an input coupler prism is varied and the angles for which a propagating optical mode is introduced into the waveguide are noted. The propagation of optical energy in the waveguide can be observed by merely placing a photodetector at the

Fig. 2.6 The prism coupler used as a device for modal analysis



output end of the waveguide. One can then determine which modes the waveguide is capable of supporting by calculating from the angle of incidence data.

An alternative method uses the prism as an output coupler. In this case, monochromatic light is introduced into the waveguide in a manner so as to excite all of the waveguide modes. For example, a diverging laser beam, either from a semiconductor laser, or from a gas laser beam passed through a lens to produce divergence, is focused onto the input face of the waveguide. Since the light is not collimated, but rather enters the waveguide at a variety of angles, some energy is

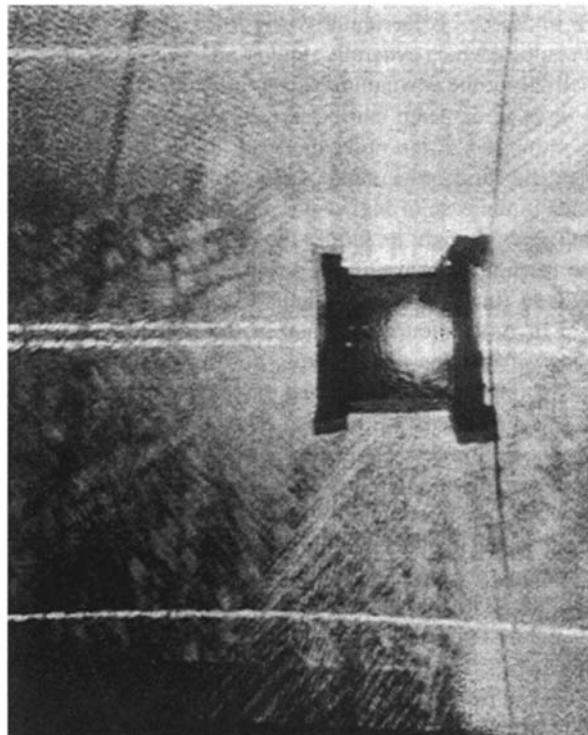


Fig. 2.7 Photograph of "m" lines produced by prism coupling of light out of a planar wave-guide. (Photo courtesy of U.S. Army ARRADCOM, Dover, NJ)

introduced into all of the waveguide modes for which the waveguide is above cutoff at the particular wavelength used. If a prism is then used as an output coupler, light from each mode emerges from the prism at a different angle. Again, the particular modes involved can be determined by calculation from the emergence angle data. Since the thickness of the waveguide is much less than its width, the emerging light from each mode appears as a band, producing a series of so-called “ m ” lines as shown in Fig. 2.7, corresponding to the particular mode number.

When the prism coupler is used to analyze the modes of a waveguide, the actual mode shape, or profile, cannot be determined in the same way as that of the scanning mirror approach of Fig. 2.3. However, the prism coupler method lets one determine how many modes can be supported by a multimode waveguide, and, as will be seen in Chap. 6, the phase velocity (hence the effective index of refraction) for each mode can be calculated from incidence and emergence angle data.

2.2 The Ray-Optic Approach to Optical Mode Theory

In Section 2.1, we considered the propagation of light in a waveguide as an electromagnetic field which mathematically represented a solution of Maxwell's wave equation, subject to certain boundary conditions at the interfaces between planes of different indices of refraction. Plane waves propagating along the z direction, supported one or more optical modes. The light propagating in each mode traveled in the z direction with a different phase velocity, which is characteristic of that mode. This description of wave propagation is generally called the physical-optic approach. An alternative method, the so-called ray-optic approach [6, 15, 16, 17], is also possible but provides a less complete description. In this latter formulation, the light propagating in the z direction is considered to be composed of plane waves moving in zig-zag paths in the x - z plane undergoing total internal reflection at the interfaces bounding the waveguide. The plane waves comprising each mode travel with the same phase velocity. However, the angle of reflection in the zigzag path is different for each mode, making the z component of the phase velocity different. The plane waves are generally represented by rays drawn normal to the planes of constant phase as shown in Fig. 2.8, which explains the name *ray-optic*.

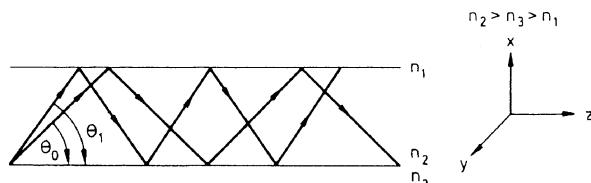


Fig. 2.8 Optical ray pattern within a multimode planar waveguide

2.2.1 Ray Patterns in the Three-Layer Planar Waveguide

The ray patterns shown in Fig. 2.8 correspond to two modes, say the TE₀ and TE₁, propagating in a three layer waveguide with $n_2 > n_3 > n_1$. The electric (\mathbf{E}) and magnetic (\mathbf{H}) fields of these plane waves traveling along zig-zag paths would add vectorially to give the \mathbf{E} and \mathbf{H} distributions of the waves comprising the same two modes, propagating in the z direction, that were described by the physical-optic model of Section 2.1. Both the ray-optic and physical-optic formulations can be used to represent either TE waves, with components E_y , H_z , and H_x , or TM waves, with components H_y , E_z and E_x .

The correlation between the physical-optic and ray-optic approaches can be seen by referring back to (2.5). The solution to this equation in the waveguiding Region 2 has the form [2.5]:

$$E_y(x, z) \propto \sin(hx + \gamma), \quad (2.7)$$

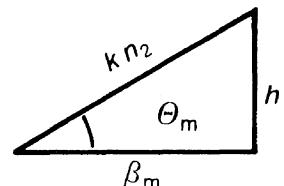
where a TE mode has been assumed, and where h and γ are dependent on the particular waveguide structure. Substituting (2.7) into (2.5) for Region 2, one obtains the condition

$$\beta^2 + h^2 = k^2 n_2^2. \quad (2.8)$$

Remembering that $k \equiv \omega/c$, it can be seen that β , h and kn_2 are all propagation constants, having units of (length)⁻¹. A mode with a z direction propagation constant β_m and an x direction propagation constant h can thus be represented by a plane wave travelling at an angle $\theta_m = \tan^{-1}(h/\beta_m)$ with respect to the z direction, having a propagation constant kn_2 , as diagrammed in Fig. 2.9. Since the frequency is constant, $kn_2 \equiv (\omega/c)n_2$ is also constant, while θ_m , β_m and h are all parameters associated with the m th mode, with different values for different modes.

To explain the waveguiding of light in a planar three-layer guide like that of Fig. 2.8 by the ray-optic method, one needs only Snell's law of refraction, coupled with the phenomenon of total internal reflection. For a thorough discussion of these basic concepts of optics see, for example, Condon [18], or Billings [19], or Bennett [20]. Consider a ray of light propagating within a three-layer waveguide structure as shown in Fig. 2.10. The light rays of Fig. 2.10a,b and c correspond to a radiation mode, a substrate mode, and a guided mode, respectively. The angles of incidence

Fig. 2.9 Geometric (vectorial) relationship between the propagation constants of an optical waveguide



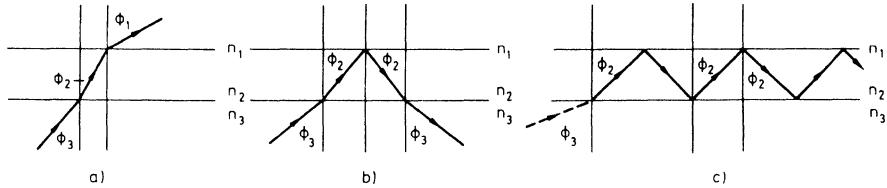


Fig. 2.10 a, b, c. Optical ray patterns for **a** air radiation modes; **b** substrate radiation modes; **c** guided mode. In each case a portion of the incident light is reflected back into layer 3; however, that ray has been omitted from the diagrams

and refraction, φ_i , with $i = 1, 2, 3$, are measured with respect to the normals to the interface planes, as is common practice in optics. From Snell's law

$$\sin \varphi_1 / \sin \varphi_2 = n_2 / n_1 \quad (2.9)$$

and

$$\sin \varphi_2 / \sin \varphi_3 = n_3 / n_2. \quad (2.10)$$

Beginning with very small angles of incidence, φ_3 , near zero, and gradually increasing φ_3 , we find the following behavior. When φ_3 is small, the light ray passes freely through both interfaces, suffering only refraction, as in Fig. 2.10a. This case corresponds to the radiation modes discussed in Section 2.1. As φ_3 is increased beyond the point at which φ_2 exceeds the critical angle for total internal reflection at the $n_2 - n_1$ interface, that light wave becomes partially confined as shown in Fig. 2.10b, corresponding to a substrate radiation mode. The condition for total internal reflection at the $n_2 - n_1$ interface is given by [19]

$$\varphi_2 \geq \sin^{-1}(n_1/n_2), \quad (2.11)$$

or, combining (2.11) and (2.10),

$$\varphi_3 \geq \sin^{-1}(n_1/n_3). \quad (2.12)$$

As φ_3 is further increased beyond the point at which φ_2 also exceeds the critical angle for total internal reflection at the $n_2 - n_3$ interface, the lightwave becomes totally confined, as shown in Fig. 2.10c, corresponding to a guided mode. In this case, the critical angle is given by

$$\varphi_2 \geq \sin^{-1}(n_3/n_2), \quad (2.13)$$

or, combining (2.2.7) and (2.2.4),

$$\varphi_3 \geq \sin^{-1}(1) = 90^\circ. \quad (2.14)$$

The conditions given by (2.11) and (2.13) for determining what type of modes can be supported by a particular waveguide as a function of φ_2 are exactly equivalent to the conditions given by (2.11) as a function of β . For example, (2.5) indicates that only radiation modes result for β less than kn_1 . Referring to Fig. 2.9, note that,

$$\varphi_2 = \beta/kn_2. \quad (2.15)$$

Thus, if $\beta \leq kn_1$,

$$\varphi_2 \leq kn_1/kn_2 = n_1/n_2, \quad (2.16)$$

which is the same condition given by (2.11). Similarly, if β is greater than kn_1 but less than kn_3 , (2.5) indicates that substrate radiation modes will be supported. Only when $\beta \geq kn_3$, can confined waveguide modes occur. From Fig. 2.9, if $\beta \geq kn_3$,

$$\sin \varphi_2 = \beta/kn_2 \geq kn_3/kn_2 \geq n_3/n_2. \quad (2.17)$$

Equation (2.17), obtained from physical-optic theory, is merely a repeat of (2.13) that resulted from the ray-optic approach. Finally, if β is greater than kn_2 ,

$$\sin \varphi_2 = \beta/kn_2 \geq 1. \quad (2.18)$$

Equation (2.18) is, of course, a physically unrealizable equality, corresponding to the physically unrealizable “a” type of modes of Fig. 2.2. Thus an equivalence has been demonstrated between the ray-optic and physical-optic approaches in regard to the determination of mode type.

2.2.2 The Discrete Nature of the Propagation Constant β

The correspondence between the ray-optic and physical optic formalisms extends beyond merely determining what type modes can be supported. It has been mentioned previously, and will be demonstrated mathematically in Chapter 3, that the solution of Maxwell's equation subject to the appropriate boundary conditions requires that only certain discrete values of β are allowed. Thus, there are only a limited number of guided modes that can exist when β is in the range

$$kn_3 \leq \beta \leq kn_2. \quad (2.19)$$

This limitation on β can be visualized quite conveniently using the ray-optic approach. The plane wavefronts that are normal to the zig-zag rays of Fig. 2.8 are assumed to be infinite, or at least larger than the cross section of the waveguide that is intercepted; otherwise they would not fit the definition of a plane wave, which requires a constant phase over the plane. Thus, there is much overlapping of the waves as they travel in the zig-zag path. To avoid decay of optical energy due to

destructive interference as the waves travel through the guide, the total phase change for a point on a wavefront that travels from the $n_2 - n_3$ interface to the $n_2 - n_1$ interface and back again must be a multiple of 2π . This leads to the condition,

$$2kn_2 t \sin \theta_m - 2\varphi_{23} - 2\varphi_{21} = 2m\pi, \quad (2.20)$$

where t is the thickness of the waveguiding Region 2, θ_m is the angle of reflection with respect to the z direction, as shown in Fig. 2.8, m is the mode number, and φ_{23} and φ_{21} , are the phase changes suffered upon total internal reflection at the interfaces. The phases $-2\varphi_{23}$ and $-2\varphi_{21}$, represent the Goos-Hänchen shifts [21, 22]. These phase shifts can be interpreted as penetration of the zig-zag ray (for a certain depth δ) into the confining layers 1 and 3 before it is reflected [6, pp. 25–29].

The values of φ_{23} and φ_{21} can be calculated from [22]:

$$\begin{aligned} \tan \varphi_{23} &= (n_2^2 \sin^2 \varphi_2 - n_3^2)^{1/2} / (n_2 \cos \varphi_2) \\ \tan \varphi_{21} &= (n_2^2 \sin^2 \varphi_2 - n_1^2)^{1/2} / (n_2 \cos \varphi_2) \end{aligned} \quad (2.21)$$

for TE waves, and

$$\begin{aligned} \tan \varphi_{23} &= n_2^2 (n_2^2 \sin^2 \varphi_2 - n_3^2)^{1/2} / (n_3^2 n_2 \cos \varphi_2) \\ \tan \varphi_{21} &= n_2^2 (n_2^2 \sin^2 \varphi_2 - n_1^2)^{1/2} / (n_1^2 n_2 \cos \varphi_2) \end{aligned} \quad (2.22)$$

for TM waves.

It can be seen that substitution of either (2.21) or (2.22) into (2.20) results in a transcendental equation in only one variable, θ_m , or φ_m , where

$$\varphi_m = \frac{\pi}{2} - \theta_m. \quad (2.23)$$

For a given m , the parameters n_1 , n_2 , n_3 and t , φ_m (or θ_m) can be calculated. Thus a discrete set of reflection angles φ_m are obtained corresponding to the various modes. However, valid solutions do not exist for all values of m . There is a cutoff condition on allowed values of m for each set of n_1 , n_2 , n_3 and t , corresponding to the point at which φ_m becomes less than the critical angle for total internal reflection at either the $n_2 - n_3$ or the $n_2 - n_1$ interface, as discussed in Section 2.2.1.

For each allowed mode, there is a corresponding propagation constant β_m given by

$$\beta_m = kn_2 \sin \varphi_m = kn_2 \cos \theta_m. \quad (2.24)$$

The velocity of the light parallel to the waveguide is then given by

$$v = c(k/\beta), \quad (2.25)$$

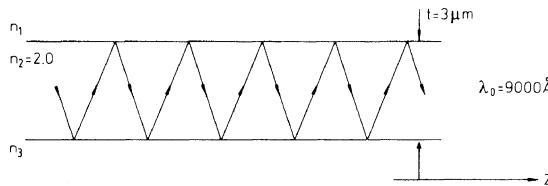
and one can define an effective index of refraction for the guide as

$$n_{\text{eff}} = c/v = \beta/k. \quad (2.26)$$

Chapter 2 has described the optical modes that can exist in a three-layer planar waveguide. We have seen that the modes can be described either by a physical-optic method, based on a solution of Maxwell's wave equation, or by a ray-optic method, relying on geometrical ray tracing principles of classical optics. In Chapter 3, the mathematical model underlying the mode theory will be developed in greater detail.

Problems

- 2.1 We wish to fabricate a planar waveguide in GaAs for light of wavelength $\lambda_0 = 1.1 \mu\text{m}$ that will operate in the single (fundamental) mode. If we assume a planar waveguide like that of Fig. 2.1 with the condition $n_2 - n_1 \gg n_2 - n_3$, what range of values can $n_2 - n_3$ have if $n_2 = 3.4$ and the thickness of the waveguiding layer $t = 3 \mu\text{m}$?
- 2.2 Repeat Problem 2.1 for the case $\lambda_0 = 1.06 \mu\text{m}$, all other parameters remaining unchanged.
- 2.3 Repeat Problems 2.1 and 2.2 for a waveguide of thickness $t = 6 \mu\text{m}$.
- 2.4 In a planar waveguide like that of Fig. 2.8 with $n_2 = 2.0$, $n_3 = 1.6$, and $n_1 = 1$, what is the angle of propagation of the lowest order mode (θ_0) when cutoff occurs? Is this a maximum or a minimum angle for θ_0 ?
- 2.5 Sketch the three lowest order modes in a planar waveguide like that of Fig. 2.8 with $n_1 = n_3 < n_2$.
- 2.6 A mode is propagating in a planar waveguide as shown with $\beta_m = 0.8 kn_2$. How many reflections at the $n_1 - n_2$ interface does the ray experience in traveling a distance of 1 cm in the z direction?



- 2.7 Show that the Goos–Hänchen phase shift goes to zero as the cutoff angle is approached for a waveguided optical mode.
- 2.8 Calculate the Goos–Hänchen shifts for a TE mode guided with $\beta = 1.85 k$ in a guide like that of Fig. 2.8, with $n_1 = 1.0$, $n_2 = 2.0$, $n_3 = 1.7$.
- 2.9 Show by drawing the vectorial relationship between the propagation constants (as in Fig. 2.9) how β , kn_2 and h change in relative magnitude and angle as one goes from the lowest-order mode in a waveguide progressively to higher-order modes.

- 2.10 A planar asymmetric waveguide is fabricated by depositing a $2\text{ }\mu\text{m}$ thick layer of Ta_2O_5 ($n = 2.09$) on to a quartz substrate ($n = 1.05$).
- How many modes can this waveguide support for light of 6328 \AA (vacuum wavelength)?
 - If a $20\text{ }\mu\text{m}$ layer of quartz ($n = 1.05$) is deposited on top of the Ta_2O_5 waveguide, how many modes can it support for light of 6328 \AA (vacuum wavelength)?
- 2.11 (a) Find the minimum required thickness for a planar slab waveguide with index of refraction = 3.5 on a substrate with index = 3.38 if it is to support the propagation of the lowest order TE mode of light with a vacuum wavelength 880 nm. The medium surrounding the waveguide and substrate is air.
- (b) If the thickness of the waveguide were increased above its minimum value by a factor of 2, and all other parameters remained unchanged, how many TE modes could be supported?

References

- A. McWhorter: Solid State Electron. **6**, 417 (1963)
- J. McKenna: Bell Syst. Techn. J. **46**, 1491 (1967)
- P.K. Tien: Appl. Opt. **10**, 2395 (1971)
- D. Marcuse: *Theory of Dielectric Optical Waveguides* (Academic, New York 1974)
- H.F. Taylor, A. Yariv: IEEE Proc. **62**, 1044 (1974)
- H. Kogelnik: Theory of dielectric waveguides, in *Integrated Optics*, T. Tamir (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979) Chap. 2
- A. Yariv: *Optical Electronics in Modern Communications*, 5th edn. (Oxford University Press, New York, Oxford 1997) Chap. 13
- D.T.F. Marple: J. Appl. Phys. **35**, 1241 (1964)
- E. Garmire, H. Stoll, A. Yariv, R.G. Hunsperger: Appl. Phys. Lett. **21**, 87 (1972)
- J. Goell: Bell Syst. Tech. J. **48**, 2133 (1969)
- P.K. Tien, G. Smolinsky, R.J. Martin: Appl. Opt. **11**, 637 (1972)
- D.H. Hensler, J. Cuthbert, R.J. Martin, P.K. Tien: Appl. Opt. **10**, 1037 (1971)
- R.G. Hunsperger, A. Yariv, A. Lee: Appl. Opt. **16**, 1026 (1977)
- Y. Luo, D.C. Hall, L. Kou, O. Blum, H. Hou, L. Steingart, J.H. Jackson: Optical Properties of $\text{Al}_x\text{Ga}_{(1-x)}\text{As}$ heterostructure native oxide planar waveguides. LEOS'99, IEEE Lasers and Electro-Optics Society 12th Annual Meeting, Orlando, Florida (1999)
- R. Ulrich, R.J. Martin: Appl. Opt. **10**, 2077 (1971)
- S.J. Maurer, L.B. Felsen: IEEE Proc. **55**, 1718 (1967)
- H.K.V. Lotsch: Optik **27**, 239 (1968)
- E.U. Condon: Electromagnetic waves, in *Handbook of Physics*, (ed.) E.U. Condon H. Odishaw (eds.) (McGraw-Hill, New York 1967) pp. 6–8
- B.H. Billings: Optics, in *American Institute of Physics Handbook*, D.E. Gray, 3rd edn. (McGraw-Hill, New York 1972) pp. 6–9
- H.E. Bennett: Reflection, in *The Encyclopedia of Physics*, (ed.) R.M. Besancon 3rd edn. (Van Nostrand Reinhold, New York 1990) pp. 1050–51
- H.K.V. Lotsch: Optik **32**, 116–137, 189–204, 299–319, 553–569 (1970/71)
- M. Born, E. Wolf: *Principles of Optics*, 3rd edn. (Pergamon, New York 1970) p. 49

Chapter 3

Theory of Optical Waveguides

Chapter 2 has reviewed the key results of waveguide theory, particularly with respect to the various optical modes that can exist in the waveguide. A comparison has been made between the physical-optic approach and the ray-optic approach in describing light propagation in a waveguide. In this chapter, the electromagnetic wave theory of the physical-optic approach is developed in detail. Emphasis is placed on the two basic waveguide geometries that are used most often in optical integrated circuits, the planar waveguide and the rectangular waveguide.

3.1 Planar Waveguides

As was mentioned previously, the planar waveguide has a fundamental geometry that has been considered by many authors in (Chapter 2 [1,2,3,4,5,6,7] as the basis for more sophisticated waveguide structure. In Sect. 3.1.1 we will follow, for the most part, the development of Taylor and Yariv (Chapter 2 [5]).

3.1.1 The Basic Three-Layer Planar Waveguide

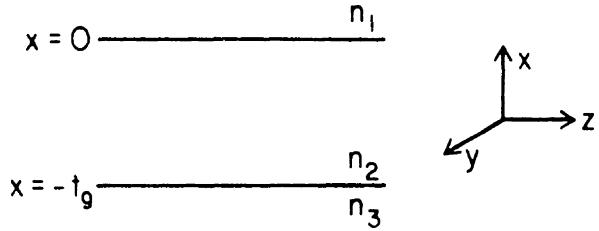
Consider the basic three-layer waveguide structure shown in Fig. 3.1. The light confining layers, with indices of refraction n_1 and n_3 , are assumed to extend to infinity in the $+x$ and $-x$ directions, respectively. The major significance of this assumption is that there are no reflections in the x direction to be concerned with, except for those occurring at the $n_1 - n_2$ and $n_2 - n_3$ interfaces. For the case of TE plane waves traveling in the z direction, with propagation constant β , Maxwell's wave equation (2.1) reduces to

$$\nabla^2 E_y = \frac{n_i^2}{c^2} \frac{\partial^2 E_y}{\partial t^2} \quad i = 1, 2, 3, \dots \quad (3.1)$$

with solutions of the form

$$E_y(x, z, t) = \mathcal{E}_y(x) e^{i(\omega t - \beta z)}. \quad (3.2)$$

Fig. 3.1 Basic three-layer planar waveguide structure



The subscript i in (3.1), of course, corresponds to a particular one of the three layers of the waveguide structure. For TE waves, it will be recalled that E_x and E_z are zero. Note also in (3.2) that \mathcal{E}_y has no y or z dependence because the planar layers are assumed to be infinite in these directions, precluding the possibility of reflections and resultant standing waves.

The transverse function $\mathcal{E}_y(x)$ has the general form

$$\mathcal{E}_y(x) = \begin{cases} = A \exp(-qx) & 0 \leq x \leq \infty \\ = B \cos(hx) + C \sin(hx) & -t_g \leq x \leq 0 \\ = D \exp[p(x + t_g)] & -\infty \leq x \leq -t_g \end{cases} \quad (3.3)$$

where A , B , C , D , q , h , and p are all constants that can be determined by matching the boundary conditions, which requires the continuity [1] of \mathcal{E}_y and $\mathcal{H}_z = (i/\omega\mu)\partial\mathcal{E}_y/\partial x$. Since the permeability μ and frequency ω are assumed to be constant, the second condition translates into a requirement that $\partial\mathcal{E}_y/\partial x$ be continuous. The constants A , B , C and D can thus be determined by making \mathcal{E}_y and $\partial\mathcal{E}_y/\partial x$ continuous at the boundary between Region 1 and Region 2 ($x = 0$), and \mathcal{E}_y continuous at $x = -t_g$. The procedure provides three equations in four unknowns, so that the solution for \mathcal{E}_y can be expressed in terms of a single constant C'

$$\mathcal{E}_y = \begin{cases} = C' \exp(-qx), & (0 \leq x \leq \infty); \\ = C' [\cos(hx) - (q/h) \sin(hx)], & (-t_g \leq x \leq 0); \\ = C'[\cos(ht_g) + (q/h) \sin(ht_g)] \exp[p(x + t_g)], & (-\infty \leq x \leq -t_g); \end{cases} \quad (3.4)$$

To determine q , h , and p , substitute (3.4) into (3.2), using the resulting expression for $E_y(x, z, t)$ in (3.1) for each of the three regions obtaining

$$\begin{aligned} q &= (\beta^2 - n_1^2 k^2)^{1/2}, \\ h &= (n_2^2 k^2 - \beta^2)^{1/2}, \\ p &= (\beta^2 - n_3^2)^{1/2}, \\ k &\equiv \omega/c. \end{aligned} \quad (3.5)$$

Note in (3.5) that q , h and p are all given in terms of the single unknown β , which is the propagation constant in the z direction. By making $\partial E_y/\partial x$ continuous at $x = -t_g$, as required, a condition on β is derived. Taking $\partial E_y/\partial x$ from (3.4) and

making it continuous at $x = -t_g$ yields the condition

$$-h \sin(-ht_g) - h(q/h) \cos(-ht_g) = p[\cos(ht_g) + (q/h) \sin(ht_g)] \quad (3.6)$$

or, after simplification,

$$\tan(ht_g) = \frac{p+q}{h(1-pq/h^2)}. \quad (3.7)$$

The transcendental equation (3.7), in conjunction with (3.5), can be solved either graphically, by plotting right and left hand sides as a function of β and noting the intersection points, or numerically on a computer. Regardless of the method of solution, the result is a set of discrete allowed values of β , corresponding to the allowed modes. For each β_m , the corresponding values of q_m , h_m and p_m can be determined from (3.5).

The one remaining unknown constant C' in (3.4) is arbitrary. However, it is convenient to normalize so that $\mathcal{E}_y(x)$ represents a power flow of one Watt per unit width in the y direction. Thus, a mode for which $E_y = A\mathcal{E}_y(x)$ has a power flow of $|A|^2$ W/m. In this case, the normalization condition is [2]

$$-\frac{1}{2} \int_{-\infty}^{\infty} E_y H_x^* dx = \frac{\beta_m}{2\omega\mu} \int_{-\infty}^{\infty} [\mathcal{E}_y^{(m)}(x)]^2 dx = 1. \quad (3.8)$$

Substituting (3.4) into (3.8) yields

$$C'_m = 2h_m \left[\frac{\omega\mu}{|\beta_m|(t_g + 1/q_m + 1/p_m)(h_m^2 + q_m^2)} \right]^{1/2}. \quad (3.9)$$

For orthogonal modes

$$\int_{-\infty}^{\infty} \mathcal{E}_y^{(l)} \mathcal{E}_y^{(m)} dx = \frac{2\omega\mu}{\beta_m} \delta_{l,m}. \quad (3.10)$$

For the case of TM modes, the development exactly parallels that which has just been performed for the TE case, except that the non-zero components are H_y , E_x , and E_z rather than E_y , H_x and H_z . The resulting field components are

$$H_y(x, z, t) = \mathcal{H}_y(x) e^{i(\omega t - \beta z)}, \quad (3.11)$$

$$E_x(x, z, t) = \frac{i}{\omega\epsilon} \frac{\partial H_y}{\partial z} = \frac{\beta}{\omega\epsilon} \mathcal{H}_y(x) e^{i(\omega t - \beta z)}, \quad (3.12)$$

$$E_z(x, z, t) = -\frac{i}{\omega\epsilon} \frac{\partial H_y}{\partial x}. \quad (3.13)$$

The transverse magnetic component $\mathcal{H}_y(x)$ is given by

$$\mathcal{H}_y(x) = \begin{cases} -C' \left[\frac{h}{\bar{q}} \cos(ht_g) + \sin(ht_g) \right] \exp[p(x + t_g)], & (-\infty \leq x \leq -t_g); \\ C' \left[-\frac{h}{\bar{q}} \cos(hx) + \sin(hx) \right], & (-t_g \leq x \leq 0); \\ C' - \frac{h}{\bar{q}} \exp[-qx], & (0 \leq x \leq \infty). \end{cases} \quad (3.14)$$

where h , q and p are again defined by (3.5), and where

$$\bar{q} = \frac{n_2^2}{n_1^2} q. \quad (3.15)$$

When boundary conditions are matched in a manner that is analogous to the TE case, it is found that only those values of β are allowed for which

$$\tan(ht_g) = \frac{h(\bar{p} + \bar{q})}{h^2 - \bar{p}\bar{q}}, \quad (3.16)$$

where

$$\bar{p} = \frac{n_2^2}{n_1^2} p. \quad (3.17)$$

The constant C' in (3.14) can be normalized so that the field represented by (3.11), (3.12), (3.13), and (3.14) carries one Watt per unit width in the y direction, leading to (Chapter 2 [5])

$$C'_m = 2 \sqrt{\frac{\omega \epsilon_0}{\beta_m t'_g}}, \quad (3.18)$$

where

$$t'_g \equiv \frac{\bar{q}^2 + h^2}{\bar{q}^2} \left(\frac{t_g}{n_2^2} + \frac{q^2 + h^2}{\bar{q}^2 + h^2} \frac{1}{n_1^2 q} + \frac{p^2 + h^2}{\bar{p}^2 + h^2} \frac{1}{n_3^2 p} \right). \quad (3.19)$$

3.1.2 The Symmetric Waveguide

A special case of the basic three-layer planar waveguide that is of particular interest occurs when n_1 equals n_3 . Such symmetric waveguides are frequently used in optical integrated circuits, for example, when a waveguiding layer with index n_2 is bounded on both surfaces by identical layers with somewhat lesser index n_1 . Multi-layer GaAlAs QIC's often utilize this type of waveguide. The equations developed

in Section 3.1.1 apply to this type of waveguide, but a major simplification is possible in the determination of which modes may be supported. In many cases, it is not required to know the β 's for the various modes. The only question being whether the waveguide is capable of guiding a particular mode or not.

A closed-form expression for the cutoff condition for TE modes can be derived in this case by referring to (2.5) and noting that, at cutoff (the point at which the field becomes oscillatory in Regions 1 and 3), the magnitude of β is given by

$$\beta = kn_1 = kn_3. \quad (3.20)$$

Substituting (3.20) into (3.5) we find that

$$\begin{aligned} p &= q = 0 \quad \text{and} \\ h &= k(n_2^2 - n_1^2)^{1/2} = k(n_2^2 - n_3^2)^{1/2}. \end{aligned} \quad (3.21)$$

Substituting (3.21) into (3.7) yields the condition

$$\tan(ht_g) = 0 \quad (3.22)$$

or

$$ht_g = m_s\pi, \quad m_s = 0, 1, 2, 3, \dots \quad (3.23)$$

Combining (3.21) and (3.23) yields

$$k(n_2^2 - n_1^2)^{1/2}t_g = m_s\pi. \quad (3.24)$$

Thus, for waveguiding of a given mode to occur, one must have

$$\Delta n = (n_2 - n_1) > \frac{m_s^2 \lambda_0^2}{4t_g^2(n_2 + n_1)}, \quad m_s = 0, 1, 2, 3, \dots, \quad (3.25)$$

where $k \equiv \omega/c = 2\pi/\lambda_0$ has been used. The cutoff condition given in (3.25) determines which modes can be supported by a waveguide with a given Δn and ratio of λ_0/t_g . It is interesting to note that the lowest-order mode ($m_s = 0$) of the symmetric waveguide is unusual in that it does not exhibit a cutoff as all other modes do. In principle, any wavelength could be guided in this mode even with an incrementally small Δn . However, for small Δn and/or large λ_0/t_g confinement would be poor, with relatively large evanescent tails of the mode extending into the substrate.

If $n_2 \cong n_1$, the cutoff condition (3.25) becomes

$$\Delta n = (n_2 - n_1) > \frac{m_s^2 \lambda_0^2}{8t_g^2 n_2}, \quad m_s = 0, 1, 2, 3, \dots \quad (3.26)$$

or, if $n_2 \gg n_1$, it is given by

$$\Delta n = (n_2 - n_1) > \frac{m_s^2 \lambda_0^2}{4t_g^2(n_2)}, \quad m_s = 0, 1, 2, 3, \dots \quad (3.27)$$

3.1.3 The Asymmetric Waveguide

Another important special case of the three-layer planar waveguide is the asymmetric waveguide, in which $n_3 \gg n_1$. Of course, n_2 must still be greater than n_3 if waveguiding is to occur. The asymmetric waveguide is often found, for example, in optical integrated circuits in which a thin film waveguide is deposited or otherwise formed on a substrate of somewhat smaller index, while the top surface of the waveguiding layer is either left open to the air or, perhaps, coated with a metal layer electrode. It is possible to derive for the case of the asymmetric guide an approximate closed form expression for the cutoff condition by using a geometrical argument comparing it to the symmetric guide [3].

Consider an asymmetric waveguide, as shown in Fig. 3.2, that has a thickness t_g equal to half the thickness of a corresponding symmetric waveguide. The two lowest order TE modes of the symmetric guide ($m_s = 0, 1$) and of the asymmetric guide ($m_a = 0, 1$) are both shown in the figure. Note that, for well-confined modes, the lower half of the ($m_s = 1$) mode of the symmetric guide corresponds closely to the ($m_a = 0$) mode of the asymmetric guide of half thickness. This fact can be used as a mathematical device to permit one to obtain a closed form expression for the cutoff condition in the case of the asymmetric waveguide.

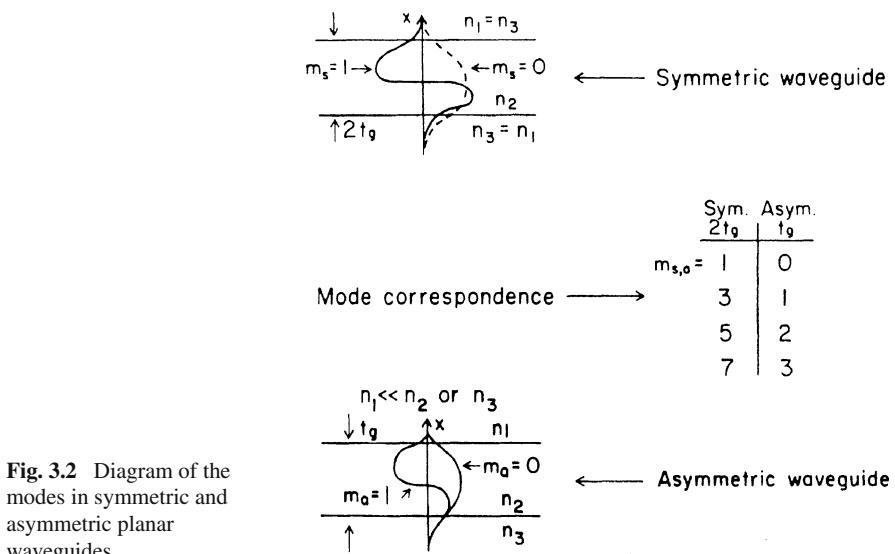


Fig. 3.2 Diagram of the modes in symmetric and asymmetric planar waveguides

Solving (3.7) for the case of the symmetric waveguide with thickness equal to $2t_g$ in the same fashion as in Section 3.1.2 yields the condition

$$\Delta n = n_2 - n_3 > \frac{m_s^2 \lambda_0^2}{4(n_2 + n_3)(2t_g)^2}, \quad m_s = 0, 1, 2, 3, \dots \quad (3.28)$$

However, the asymmetric waveguide supports only modes corresponding to the odd modes of a symmetric guide of twice its thickness. Hence, the cutoff condition for the asymmetric guide is given by

$$\Delta n = n_2 - n_3 > \frac{m_a^2 \lambda_0^2}{16(n_2 + n_3)t_g^2}, \quad (3.29)$$

where m_a are the elements of the subset consisting of odd values of m_s . This can be conveniently expressed by

$$m_a = (2m + 1), \quad m = 0, 1, 2, 3, \dots \quad (3.30)$$

Assuming that $n_2 \cong n_3$ (3.29) becomes

$$\Delta n = n_2 - n_3 > \frac{(2m + 1)^2 \lambda_0^2}{32n_2 t_g^2}, \quad m = 0, 1, 2, 3, \dots \quad (3.31)$$

While the cutoff conditions (3.31) and (3.25) are valid only for the special cases defined, they offer a convenient means to estimate how many modes can be supported by a particular waveguide. To answer this question in the general case, or to determine the β_m for the various modes, one would have to solve the transcendental equation (3.7).

Although (3.31) has been derived for the case of TE waves, it can be shown that it also holds for TM waves as long as $n_2 \cong n_3$. Thus, the asymmetric waveguide is seen to have a possible cutoff for all modes, unlike the symmetric waveguide for which the TE_0 mode cannot be cut off. This makes the asymmetric waveguide particularly useful as an optical switch, as will be discussed in Chapter 9.

3.2 Rectangular Waveguides

The planar waveguides discussed in the previous section are useful in many integrated optic applications in spite of the fact that they provide confinement of the optical fields in only one dimension. Even relatively complex optical integrated circuits, such as the rf spectrum analyzer of Mergerian and Malarkey [4] and the optical spectrum analyzer of Madsen et al. [5] can be fabricated using planar waveguides. However, other applications require optical confinement in two dimensions. Use of a “stripe” geometry waveguide of rectangular cross-section can yields a laser

with reduced threshold current and single mode oscillation [6], or an electro-optic modulator with reduced drive power requirement [7]. Sometimes two-dimensional confinement is required merely to guide light from one point on the surface of an OIC to another, to interconnect two circuit elements in a manner analogous to that of the metallic stripes used in an electrical integrated circuit.

3.2.1 Channel Waveguides

The basic rectangular waveguide structure consists of a waveguide region of index n_1 surrounded on all sides by a confining medium of lesser index n_2 , as shown in Fig. 3.3. Such waveguides are often called *channel* guides, *strip* guides, or *3-dimensional* guides. It is not necessary that the index in the confining media be the same in all regions. A number of different materials, all with indices less than n_1 , may be used to surround the guide. However, in that case, the modes in the waveguide will not be exactly symmetric. The exact solution of the wave equation for this general case is extremely complicated, and has not been obtained yet.

However, Marcatill [10] has derived an approximate solution to the rectangular channel waveguide problem, by analyzing the structure shown in Fig. 3.4, which is still fairly general. The key assumption made in Marcatili's analysis is that the modes are well guided, i.e., well above cutoff, so that the field decays exponentially in Regions 2, 3, 4, and 5, with most of the power being confined to Region 1. The magnitudes of the fields in the shaded corner regions of Fig. 3.4 are small enough to be neglected. Hence, Maxwell's equations can be solved by assuming relatively simple sinusoidal and exponential field distributions, and by matching boundary conditions only along the four sides of Region 1. The waveguide is found to support a discrete number of guided modes that can be grouped into two families, E_{pq}^x and E_{pq}^y , where the mode numbers p and q correspond to the number of peaks in the field

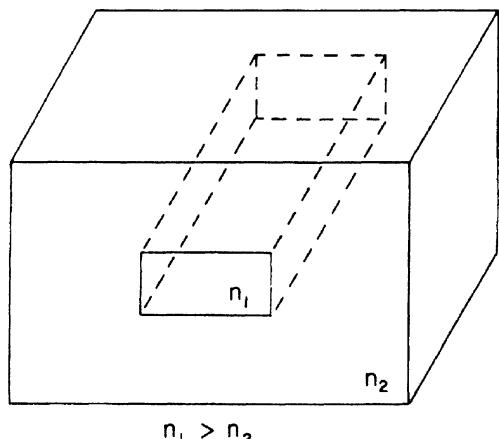
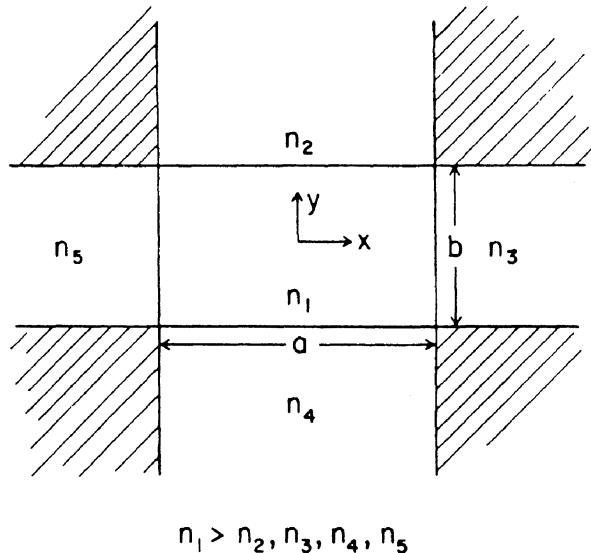


Fig. 3.3 Basic rectangular dielectric waveguide structure

Fig. 3.4 Cross-sectional view of a rectangular dielectric waveguide bounded by regions of smaller index of refraction



distribution in the x and y directions, respectively. The transverse field components of the E_{pq}^x modes are E_x and H_y , while those of the E_{pq}^y modes are E_y and H_x . The E_{11}^y (fundamental) mode is sketched in Fig. 3.5. Note that the shape of the mode is characterized by extinction coefficients η_2 , ξ_3 , η_4 , and ξ_5 in the regions where it is exponential, and by propagation constants k_x and k_y in Region 1.

Quantitative expressions for k_x , k_y , η and ξ can be determined for the E_{pq}^y modes as follows. The field components in the five regions shown in Fig. 3.4 (designated by $v = 1, 2, 3, 4, 5$) have the form:

$$H_{xv} = \exp(-ik_z z + i\omega t) \begin{cases} M_1 \cos(k_x x + \alpha) \cos(k_y y + \beta) & \text{for } v = 1 \\ M_2 \cos(k_x x + \alpha) \exp(-ik_{y2} y) & \text{for } v = 2 \\ M_3 \cos(k_y y + \beta) \exp(-ik_{x3} x) & \text{for } v = 3 \\ M_4 \cos(k_x x + \alpha) \exp(ik_{y4} y) & \text{for } v = 4 \\ M_5 \cos(k_y y + \beta) \exp(ik_{x5} x) & \text{for } v = 5. \end{cases} \quad (3.32)$$

$$H_{yv} = 0, \quad (3.33)$$

$$H_{zv} = -\frac{i}{k_z} \frac{\partial^2 H_{xv}}{\partial x \partial y}, \quad (3.34)$$

$$E_{xv} = -\frac{1}{\omega \epsilon_0 n_v^2 k_z} \frac{\partial^2 H_{xv}}{\partial x \partial y}, \quad (3.35)$$

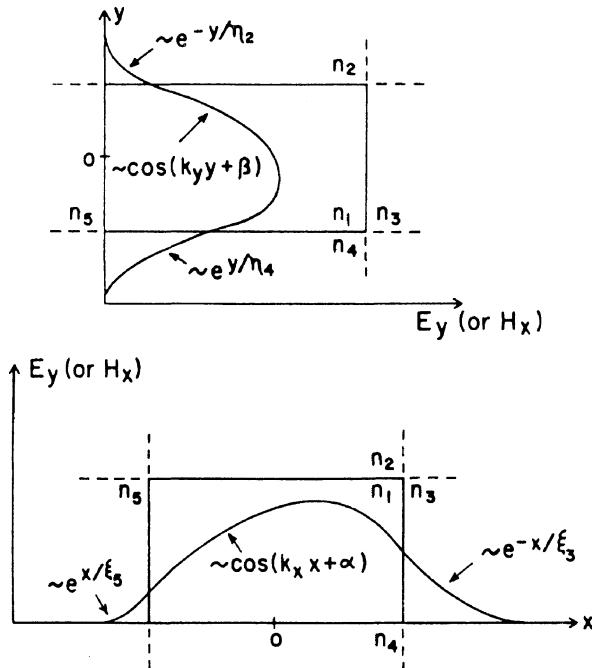


Fig. 3.5 Sketch of a typical E_{11}^y mode

$$E_{yv} = \frac{k^2 n_v^2 - k_{yv}^2}{\omega \epsilon_0 n_v^2 k_z} H_{xv}, \quad (3.36)$$

$$E_{zv} = -\frac{i}{\omega \epsilon_0 n_v^2} \frac{\partial H_{xv}}{\partial y}, \quad (3.37)$$

where M_v is an amplitude constant, ω is the angular frequency and ϵ_0 is the permittivity of free space. The phase constants, α and β locate the field maxima and minima in Region 1, and k_{xv} and k_{yv} ($v = 1, 2, 3, 4, 5$) are the transverse propagation constants along the x and y directions in the various media. Matching the boundary conditions requires the assumption that

$$k_{x1} = k_{x2} = k_{x4} = k_x, \quad (3.38)$$

and

$$k_{y1} = k_{y3} = k_{y5} = k_y. \quad (3.39)$$

Also, it can be shown that

$$k_z = (k_1^2 - k_x^2 - k_y^2)^{1/2}, \quad (3.40)$$

where

$$k_1 = kn_1 = \frac{2\pi}{\lambda_0} n_1, \quad (3.41)$$

is the propagation constant of a plane wave with free-space wavelength λ_0 in a medium of refractive index n_1 . Assuming that n_1 is only slightly larger than the other n_v , as is usually the case in an OIC, leads to the condition

$$k_x \text{ and } k_y \ll k_z \quad (3.42)$$

Note that (3.42) corresponds, in ray-optics terminology, to a grazing incidence of the ray at the surfaces of the waveguiding Region 1. Calculations show that the two significant components for the E_{pq}^y modes are H_x and E_y .

Matching field components at the boundaries of Region 1 yields the transcendental equations

$$k_x a = p\pi - \tan^{-1} k_x \xi_3 - \tan^{-1} k_x \xi_5, \quad (3.43)$$

and

$$k_y b = q\pi - \tan^{-1} \frac{n_2^2}{n_1^2} k_y \eta_2 - \tan^{-1} \frac{n_4^2}{n_1^2} k_y \eta_4, \quad (3.44)$$

where the \tan^{-1} functions are to be taken in the first quadrant, and where

$$\xi_5^3 = \frac{1}{|k_{x_5}|} = \frac{1}{\left[\left(\frac{\pi}{A_3} \right)^2 - k_x^2 \right]}, \quad (3.45)$$

$$\eta_4^2 = \frac{1}{|k_{y_4}|} = \frac{1}{\left[\left(\frac{\pi}{A_4} \right)^2 - k_y^2 \right]^{1/2}}, \quad (3.46)$$

and

$$A_v = \frac{\pi}{(k_1^2 - k_v^2)^{1/2}} = \frac{\lambda_0}{2(n_1^2 - n_v^2)^{1/2}}, \quad v = 2, 3, 4, 5. \quad (3.47)$$

The transcendental equations (3.43) and (3.44) cannot be solved exactly in closed form. However, one can assume for well confined modes that most of the power is in Region 1. Hence,

$$\left(\frac{K_x A_5^3}{\pi}\right) \ll 1 \text{ and } \left(\frac{K_y A_4^2}{\pi}\right)^2 \ll 1. \quad (3.48)$$

Using the assumptions of (3.48), approximate solutions of (3.43) and (3.48) for k_x and k_y can be obtained by expanding the \tan^{-1} functions in a power series, keeping only the first two terms. Thus,

$$k_x = \frac{p\pi}{a} \left(1 + \frac{A_3 + A_5}{\pi a}\right)^{-1} \quad (3.49)$$

$$k_y = \frac{q\pi}{b} \left(\frac{1 + n_2^2 A_2 + n_4^2 A_4}{\pi n_1^2 b}\right)^{-1} \quad (3.50)$$

Substituting (3.49) and (3.50) in (3.40), (3.45) and (3.46) expressions for k_z , ξ_3 , ξ_5 , η_2 and η_4 can be obtained.

$$k_2 = \left[k_1^2 - \left(\frac{\pi p}{a}\right)^2 \left(1 + \frac{A_3 + A_5}{\pi a}\right)^{-2} - \left(\frac{\pi q}{b}\right)^2 \left(1 + \frac{n_2^2 A_2 + n_4^2 A_4}{\pi n_1^2 b}\right)^{-2} \right]^{1/2}. \quad (3.51)$$

$$\xi_5^3 = \frac{A_5}{\pi} \left[1 - \left(\frac{p A_3}{a} \frac{1}{1 + \frac{A_3 + A_5}{\pi a}} \right)^2 \right]^{-1/2}. \quad (3.52)$$

$$\eta_2^2 = \frac{A_2}{\pi} \left[1 - \left(\frac{q A_4}{b} \frac{1}{1 + \frac{n_2^2 A_2 + n_4^2 A_4}{\pi n_1^2 b}} \right)^2 \right]^{-1/2}. \quad (3.53)$$

The E_{pq}^y modes are polarized such that E_y is the only significant component of electric field; E_x and E_z are negligibly small. It can be shown for the case of the E_{pq}^x modes that E_x is the only significant electric field component, with E_y and E_z being negligible. To develop relationships for the E_{pq}^x modes corresponding to those that have been derived from the E_{pq}^y modes, one can merely change E to H , μ_0 to $-\varepsilon_0$, and vice-versa in the various equations. As long as the assumption is made that n_1 is only slightly larger than the indices of the surrounding media, i.e.,

$$\frac{1}{n_1}(n_1 - n_v) \ll 1, \quad (3.54)$$

then k_z , ξ_3 , and η_2 are still given by (3.51), (3.52), and (3.53), respectively, for the E_{pq}^y modes just as for the E_{pq}^y modes.

Marcatilli's analysis of the rectangular three-dimensional waveguide is very useful in designing such structures, even though it features an approximate solution

to Maxwell's equations. It must be remembered that the theory assumes well confined modes. When waveguide dimensions a and b are small enough, compared to the wavelength, the theory becomes inaccurate for the mode [8]. The rectangular waveguide can also be analyzed by using the "Finite Element Method", as described by Katz [9]. Regardless of the method that is used, analysis of rectangular waveguide structures is always mathematically complicated. Fortunately, there are now a number of commercially available computer programs to do it [10].

3.2.2 Strip-Loaded Waveguides

It is possible to make a three-dimensional waveguide, in which there is confinement in both the x and y dimensions, without actually surrounding the waveguide with materials of lesser index. This is done by forming a strip of dielectric material of lesser index, n_3 , on top of a planar waveguide, with index n_1 , as shown in Fig. 3.6. Such a structure is usually called either a *strip-loaded* waveguide, or an *optical stripline*. The presence of the loading strip on top of the waveguiding layer makes the effective index in the region beneath it, n'_{eff1} larger than the effective index, n_{eff1} in the adjacent regions. Thus there can be confinement in the y direction as well as in the x direction. The physical nature of this phenomenon can be visualized best by using the ray optic approach. Consider a particular mode propagating in the z direction but consisting of a plane wave following the usual zig-zag path in the waveguiding layer (Fig. 2.10). Since n_3 is larger than n_4 , the wave penetrates slightly more at the $n_1 - n_3$ interface than it would at the $n_1 - n_4$ interface; thus,

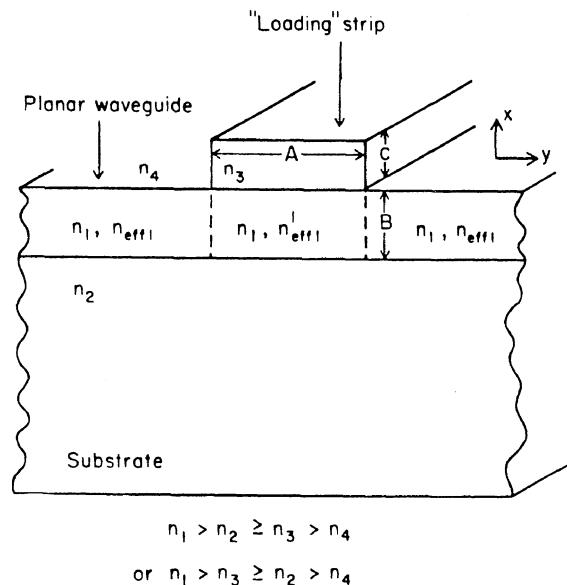
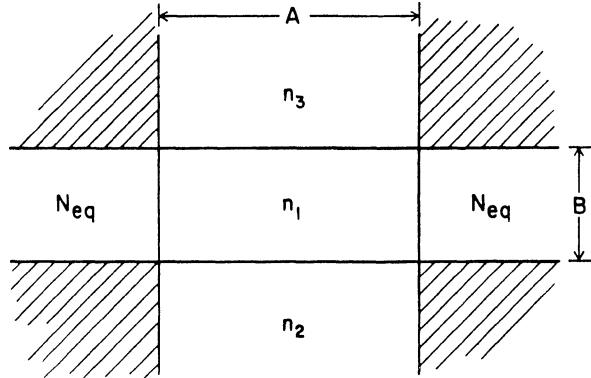


Fig. 3.6 Diagram of dielectric strip-loaded waveguide. Effective index of refraction, as well as the bulk index, are indicated in the wave-guiding layer.

$$n'_{\text{eff1}} > n_{\text{eff1}}$$

Fig. 3.7 Cross-sectional view of rectangular dielectric waveguide equivalent to the strip-loaded waveguide of Fig. 3.6



the effective height of the waveguide is greater under the loading strip than it is in the regions on either side. This means that the zig-zag path of the plane wave would be slightly longer under the loading strip, leading to the result that

$$n'_{\text{eff}1} = \frac{\beta'}{k} > n_{\text{eff}1} = \frac{\beta}{k}. \quad (3.55)$$

Furuta et al. [11] have used the effective index of refraction method to analyze a strip-loaded guide like that of Fig. 3.6, and have shown that its waveguiding properties are equivalent to those of a dielectric waveguide like that shown in Fig. 3.7, where the equivalent index in the side confining layers is given by

$$n_{\text{eq}} = \left(n_1^2 - n'_{\text{eff}1}^2 + n_{\text{eff}1}^2 \right)^{1/2}. \quad (3.56)$$

The propagation constants of the rectangular waveguide of Fig. 3.7 can be determined by using Marcatili's method, described in Sect. 3.2.1. The effectiveness of this approach has been demonstrated [11] by comparing theoretical predictions with experimental observations of waveguiding of 6328 Å light in strip-loaded waveguides.

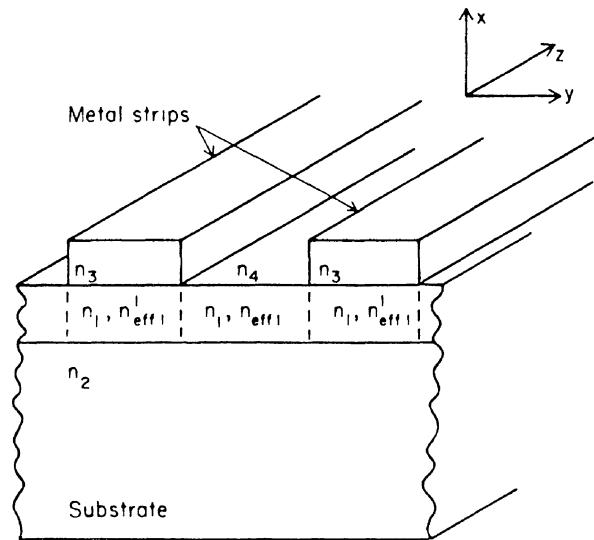
In that case, the waveguides were thin (approx. 0.6 μm) glass films ($n_1 = 1.712$), on glass substrates with $n_2 = 1.662$. The index of the loading strip was 1.592 and its cross-sectional dimensions were 0.7 μm thick × 14 μm wide. Ramaswamy [12] has also used the effective index method to analyze strip-loaded guides and found good agreement with experimental results.

A strip-loaded waveguide can also be made with the index of the loading strip, n_3 , equal to that of the guide, n_1 . That type of waveguide is usually called a *ridge*, or *rib*, guide. Kogelnik [13] has applied the effective index method to that type of guide to show that propagation is characterized by a phase constant

$$\beta = kN, \quad (3.57)$$

Fig. 3.8 Metal strip-loaded waveguide. Effective indices of refraction, as well as the bulk index, are indicated in the waveguiding layer.

$$n'_{\text{eff}1} < n_{\text{eff}1}$$



where N is given by

$$N^2 = n_{\text{eff}1}^2 + b(n'_{\text{eff}1}^2 - n_{\text{eff}1}^2). \quad (3.58)$$

In this case, $n'_{\text{eff}1}$ is the effective index in a guide having a height equal to the sum of the thicknesses of the waveguiding layer and the loading strip. The parameter b is a normalized guide index, given by

$$b = (N^2 - n_2^2) / (n_1^2 - n_2^2). \quad (3.59)$$

Metallic loading strips can also be used to produce optical striplines. In that case, two metal loading strips are placed on the surface of the waveguiding layer, on either side of the region in which confinement is desired, as shown in Fig. 3.8. Since penetration of the guided waves is deeper at the $n_1 - n_4$ interface than at the $n_1 - n_3$ interface, the desired confinement in the y dimension is obtained just as in the case of the dielectric strip-loaded waveguide. Metal strip-loaded waveguides are particularly useful in applications such as electro-optic modulators, where surface metal electrodes are desired, since these metal stripes can perform the additional function of defining the waveguide [14].

In principle, it is expected that strip-loaded waveguides should have less optical loss than rectangular channel dielectric waveguides, because scattering due to side-well roughness is reduced. Experimental results seem to support that hypothesis. Blum et al. [15] have reported a loss coefficient of 1 cm^{-1} in GaAs strip-loaded guides, while Reinhart et al. [16] have measured waveguide loss of less than 2 cm^{-1} in GaAs-GaAlAs rib waveguides. The greatest concern regarding the usefulness of strip-loaded waveguides is that the small effective index difference in the

y dimension, produced by the loading effect, will be insufficient to limit radiation loss occurring from bends in the guide. However, 90° bends with radii as small as 2.5 mm have been made [11] in glass strip-loaded guides with no observation of excessive radiation loss.

Chapter 3 has described basic theoretical models that can be used to calculate the propagation characteristics of the waveguide types that are commonly used in optical integrated circuits. Additional information on this topic is available, for example, in the books by Unger [17], Tamir [18], Fernandez [19], Okamoto [20] and Calvo [21] and in a variety of journal articles, [22–61], which have been listed by title in the references for this chapter to facilitate referral. Because of the complexity of mathematical models for optical waveguides, computer programs have been developed to do waveguide design and analysis [62–68]. In the next chapter, methods of fabricating these waveguides will be described.

Problems

- 3.1 For a basic three-layer planar waveguide such as that shown in Fig. 3.1, show that the phase constant h in Region 2, and the extinction coefficients q and p in Region 1 and 3, respectively, are given by

$$\begin{aligned} q &= (\beta^2 - n_1^2 k^2)^{1/2} \\ p &= (\beta^2 - n_3^2 k^2)^{1/2} \\ h &= (n_2^2 k^2 - \beta^2)^{1/2}, \end{aligned}$$

where β is the propagation constant in the z direction and $k \equiv \omega/c$.

- 3.2 Show that the change in index of refraction Δn required for waveguiding of the m th order mode in an asymmetric waveguide is given by

$$\Delta n = n_2 - n_3 > \frac{(2m+1)^2 \lambda_0^2}{32n_2 t_g^2}, \quad m = 0, 1, 2, 3,$$

where all terms are as defined in Section 3.1.3.

- 3.3 Sketch the cross-sectional electric field distributions of the E_{12}^y and E_{22}^y modes in a rectangular waveguide.
 3.4 Sketch the cross-sectional optical power distributions of the E_{33}^y and E_{22}^y modes in a rectangular waveguide.
 3.5 A *symmetric* planar waveguide has a waveguiding layer with thickness = 3 μm and index $n_2 = 1.5$. The waveguide is excited with a He-Ne laser ($\lambda_0 = 0.6328 \mu\text{m}$).

Assume that the surrounding medium is air (i.e., $n_1 = n_3 = 1$)

- How many TE modes can the waveguide support?
- Suppose that instead of the surrounding medium being air it is dielectric so that $n_1 = n_3 = 1.48$, how many TE modes can now propagate?

- (c) In the case of the waveguide of part (b), how thin can the waveguiding layer be made and still allow the lowest order mode to propagate?
- 3.6 A planar dielectric waveguide is fabricated by sputtering a thin glass film ($n_2 = 1.62$) on top of a glass substrate ($n_3 = 1.346$). The light source is a He-Ne laser operating at wavelength $\lambda_0 = 6328 \text{ \AA}$.
- What is the maximum thickness of the waveguiding layer for the guide to propagate only one mode?
 - If the propagation constant in the direction of propagation is $\beta = 1.53 \times 10^5 \text{ cm}^{-1}$, what is the bounce angle θ_m of the mode? (Make a sketch and label the angle for clarity.)
- 3.7 For an asymmetric planar waveguide like that shown in Fig. 3.1 with $n_1 = 1.0$, $n_2 = 1.65$, $n_3 = 1.52$, $t_g = 1.18 \mu\text{m}$ and $\lambda_0 = 0.63 \mu\text{m}$, find the number of allowed TE modes.
- 3.8 Explain why it is easier to make a single-mode waveguide in a symmetric three-layer waveguide than it is in an asymmetric three-layer waveguide.
- 3.9 (a) What is an “optical stripline”?
- Using the ray-optic approach, sketch the propagation of a lightwave in a strip-loaded waveguide like that shown in Fig. 3.6.
- 3.10 Consider a plane wave traveling in a zig-zag path along the length of a strip-loaded waveguide. Is the velocity of the wave under the loading strip greater or smaller than it would be without the loading strip?

References

- W. Hayt, Jr.: *Engineering Electromagnetics*, 4th edn. (McGraw-Hill, New York 1981) p. 151 and p. 317
- A. Yariv: *Optical Electronics*, 4th edn. (Holt, Rinehart and Winston, New York 1991) p. 490
- D. Hall, A. Yariv, E. Garmire: Opt. Commun. **1**, 403 (1970)
- D. Mergerian, E. Malarkey: Microwave J. **23**, 37 (1980)
- C.K. Madsen, J. Wagener, T.A. Strasser, D. Muhlner, M.A. Milbrodt, E.J. Laskowski, J. DeMarco: IEEE J. Selected Topics in Quantum Electronics **4**, 925 (1998)
- H. Kressel, M. Ettenberg, J. Wittke, I. Laddany: Laser diodes and LEDs for fiber optical communications, in *Semiconductor Devices*, H. Kressel, (ed.), 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982) pp. 23–25
- S. Somekh, E. Germaire, A. Yariv, H. Garvin, R.G. Hunsperger: Appl. Opt. **13**, 327 (1974)
- A.A.J. Marcatilli: Bell Syst. Tech. J. **48**, 2071 (1969)
- J. Katz: Novel solution of 2-D waveguides using the finite element method, Appl. Opt. **21**, 2747 (1982)
- See, e.g., the following suppliers: Bay Technology (<http://www.bay-technology.com>) BBV-Software (<http://www.bbv-software.com>) Breault Research (<http://www.breault.com>) Integrated Optical Software (<http://www.ios-gmbh.de/>) Optiwave (<http://www.optiwave.com/>) RSOFT (<http://www.rsoftinc.com/home.htm>) Stellar Software (<http://www.stellarsoftware.com/>) Catalog (A listing of additional software sources.) (<http://home.earthlink.net/~skywise711/LasersOptics/Software/PhotonicSoftware.html>)
- H. Furuta, H. Noda, A. Ihaya: Appl. Opt. **13**, 322 (1974)
- V. Ramaswamy: Bell Syst. Tech. J. **53**, 697 (1974)

13. H. Kogelnik: Theory of dielectric waveguides, in *Integrated Optics*, T. Tamir (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979) Chap. 2
14. J. Campbell, F. Blum, D. Shaw, K. Shaw, K. Lawley: Appl. Phys. Lett. **27**, 202 (1975)
15. F. Blum, D. Shaw, W. Holton: Appl. Phys. Lett. **25**, 116 (1974)
16. F. Reinhart, R. Logan, T. Lee: Appl. Phys. Lett. **24**, 270 (1974)
17. H.G. Unger: *Planar Optical Waveguides and Fibers* (Oxford University Press, Oxford 1978)
18. H. Kogelnik: Theory of optical waveguides, in *Guided Wave Optoelectronics*, T. Tamir (ed.) 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990)
19. F.A. Fernandez, Y. Lu: *Microwave and Optical Waveguide Analysis by the Finite Element Method* (Research Studies Press/ Wiley, Taunton/New York, 1996)
20. K. Okamoto: *Fundamentals of Optical Waveguides*, 2nd edn. (Elsevier Academic Press, New York, 2006)
21. M.L. Calvo and V. Lakshminarayanan: *Optical Waveguides: From Theory to Applied Technologies* (CRC Press, Taylor and Francis Group, Boca Raton, FL, 2007)
22. T. Tamir: Microwave modeling of periodic waveguides. IEEE Trans. MTT-**29**, 979 (1981)
23. T. Tamir: Guided-wave methods for optical configurations. Appl. Phys. **25**, 201 (1981)
24. E.A. Kolosovsky, D.V. Petrov, A.V. Tsarey, I.B. Yakovkin: An exact method for analyzing light propagation in anisotropic inhomogeneous optical waveguides. Opt. Commun. **43**, 21 (1982)
25. K. Yasumoto, Y. Oishi: A new evaluation of the Goos Hänchen shift and associated time delay. J. Appl. Phys. **54**, 2170 (1983)
26. F.P. Payne: A new theory of rectangular optical waveguides. Opt. Quant. Electron. **14**, 525 (1982)
27. H. Yajima: Coupled-mode analysis of anisotropic dielectric planar branching waveguides. IEEE J. LT-**1**, 273 (1983)
28. S.A. Shakir, A.F. Turner: Method of poles for multiyer thin film waveguides. Appl. Phys. A **29**, 151 (1982)
29. W.H. Southwell: Ray tracing in gradient-index media. J. Opt. Soc. Am. **72**, 909 (1982)
30. J. Van Roey, J. Vander Donk, P.E. Lagasse: Beam-propagation method: Analysis and assessment. J. Opt. Soc. Am. **71**, 803 (1981)
31. J. Nezval: WKB approximation for optical modes in a periodic planar waveguide. Opt. Commun. **42**, 320 (1982)
32. Ch. Pichot: Exact numerical solution for the diffused channel waveguide. Opt. Commun. **41**, 169 (1982)
33. V. Ramaswamy, R.K. Lagu: Numerical field solution for an arbitrary asymmetrical graded-index planar waveguide. IEEE J. LT-**1**, 408 (1983)
34. Y. Li: Method of successive approximations for calculating the eigenvalues of optical thin-film waveguides. Appl. Opt. **20**, 2595 (1981)
35. J.P. Meunier, J. Piggeon, J.N. Massot: A numerical technique for the determination of propagation characteristics of inhomogeneous planar optical waveguides. Opt. Quant. Electron. **15**, 77 (1983)
36. M. Belanger, G.L. Yip: Mode conversion analysis in a single-mode planar taper optical waveguide. J. Opt. Soc. Am. **72**, 1822 (1982)
37. E. Khular, A. Kumar, A. Sharma, I.C. Goyal, A.K. Ghatak: Modes in buried planar optical waveguide with graded-index profiles. Opt. Quant. Electron. **13**, 109 (1981) A.K. Ghatak: Exact modal analysis for buried planar optical waveguides with asymmetric graded refractive index. Opt. Quant. Electron. **13**, 429 (1981)
38. A. Hardy, E. Kapon, A. Katzir: Expression for the number of guided TE modes in periodic multilayer waveguides. J. Opt. Soc. Am. **71**, 1283 (1981)
39. L. Eyges, P. Wintersteiner: Modes in an array of dielectric waveguides. J. Opt. Soc. Am. **71**, 1351 (1981) L. Eyges, P.D. Gianino: Modes of cladded guides of arbitrary cross-sectional shape. J. Opt. Soc. Am. **72**, 1606 (1982)
40. P.M. Rodhe: On radiation in the time-dependent coupled power theory for optical waveguides. Opt. Quant. Electron. **15**, 71 (1983)

41. L. McCaughan, E.E. Bergmann: Index distribution of optical waveguides from their mode profile. *IEEE J. LT-1*, 241 (1983)
42. H. Kogelnik: Devices for lightwave communications, in *Lasers and Applications*, W.O.N. Guimaraes, C.-T. Lin, A. Mooradian (eds.) Springer Ser. Opt. Sci., Vol. 26 (Springer, Berlin, Heidelberg 1981)
43. R.E. Smith, S.N. Houde-Walter, G.W. Forbes: Mode determination for planar waveguide using the four-sheeted dispersion relation. *IEEE J. Quant. Electron.* **28**, 1520 (1992)
44. H. Renner: Bending losses of coated single-mode fibers: a simple approach. *IEEE J. Lightwave Tech.* **10**, 544 (1992)
45. F. Olyslager, D. De Zutter: Rigorous boundary integral equation solution for general isotropic and uniaxial anisotropic dielectric waveguides in multilayered media including losses, gain and leakage. *IEEE Trans. Micro. Theor. Tech.* **41**, 1385 (1993)
46. J.W. Mink, F.K. Schwering: A hybrid dielectric slab-beam waveguide for the submillimeter wave region. *IEEE Trans. Micro. Theor. Tech.* **41**, 1720 (1993)
47. F. Di Pasquale, M. Zoboli, M. Federighi, I. Massarek: Finite-element modeling of silica waveguide amplifiers with high erbium concentration. *IEEE J. Quant. Electron.* **30**, 1277 (1994)
48. G.R. Hadley, R.E. Smith: Full-vector waveguide modeling using an iterative finite-difference method with transparent boundary conditions. *IEEE J. Lightwave Tech.* **13**, 465 (1995)
49. S.M. Tseng, J.H. Zhan: A new method of finding the propagation constants of guided modes in slab waveguides containing lossless and absorbing media. Proc. Lasers and Electro-Optics Society Annual Meeting, LEOS'97 (1997) pp. 512–513.
50. O. Mitomi, K. Kasaya: Wide-angle finite-element beam propagation method using Padé approximation. *Electron. Lett.* **33**, 1461 (1997)
51. K. Kawano, T. Kitoh, M. Kohtoku, T. Takeshita, Y. Hasumi: 3-D semivectorial analysis to calculate facet reflectivities of semiconductor optical waveguides based on the bi-directional method of line BPM (MoL-BPM). *IEEE Photo. Tech. Lett.* **10**, 108 (1998)
52. G. Tartarini, H. Renner: Efficient finite-element analysis of tilted open anisotropic optical channel waveguides. *IEEE Micro. Guided Wave Lett.* **9**, 389 (1999)
53. A.A. Abou El-Fadl, K.A. Mostafa, A.A. Abelenin, T.E. Taha: New technique for analysis of multimode diffused channel optical waveguides. Digest, IEEE Conf. Infrared and Millimeter Waves (2000) pp. 229–230
54. R. Scarmozzino, A. Gopinath, R. Pregla, S. Helfert: Numerical techniques for modeling guided-wave photonic devices. *IEEE J. Selected Topics in Quant. Electron.* **6**, 150 (2000)
55. K. Saitoh, M. Koshiba: Approximate scalar finite-element beam-propagation method with perfectly matched layers for anisotropic optical waveguides. *IEEE J. Lightwave Techno.* **19**, 786 (2001)
56. C.R. Doerr: Beam propagation method tailored for step-index waveguides. *IEEE Phot. Technol. Lett.* **13**, 130 (2001)
57. A. Giorgio, A.G. Perri, M.N. Armenise: Modelling waveguiding photonic bandgap structures by leaky mode propagation method. *Electron. Lett.* **37**, 835 (2001)
58. A. Sharma: Analysis of integrated optical waveguides: variational method and effective-index method with built-in perturbation correction, *J. Opt. Soc. Am. A* **18**, 1383 (2001)
59. R. Pregla: Modeling of planar waveguides with anisotropic layers of variable thickness by the method of lines, *Opt. Quant. Electron.* **35**, 533 (2003)
60. T. Miyamoto, M. Momoda, K. Yasumoto: Numerical analysis for three-dimensional optical waveguides with periodic structure using Fourier series expansion method, *Electron. Communi. Jap. (Part II: Electron.)* **86**, 22 (2003)
61. M.A. Boroujeni, M. Shahabadi: Full-wave analysis of lossy anisotropic optical waveguides using a transmission line approach based on a Fourier method, *J. Opt. A:* **8** 1080 (2006)
62. J. Costa, D. Pereira, A. Giarola: Analysis of optical waveguides using Mathematica (R) Microwave and Optoelectronics Conference, Proceedings, 1997, SBMO/IEEE MTT-S International **1**, 91 (1997)

63. Y. Moreau, J. Porque, P. Coudray, P. Etienne, K. Kribich: New simulation tools for complex multilevel optical circuits, SPIE International Conference, Optical Design and Analysis Software, Denver, CO (1999)
64. M.F. van der Vliet, G. Beelen: Design and simulation tools for integrated optic circuits, Proc. SPIE **3620**, 174 (1999)
65. M.R. Amersfoort: Design and Simulation Tools for Photonic Integrated Circuits, LEOS 2000 Annl. Meet. Conf. Proc. **2**, 774 (2000)
66. M. Amersfoort: Simulation and Design Tools Address Demands of WDM, Laser Focus World, (March 2001), pp.129–32
67. T.G. Nguyen, A. Mitchell: Analysis of optical waveguides with multilayer dielectric coatings using plane wave expansion, *J. Lightwave Techn.* **24**, 635 (2006)
68. P.R. Chaudhuri, S. Roy: Analysis of arbitrary index profile planar optical waveguides and multilayer nonlinear structures: a simple finite difference algorithm, *Opt. Quant. Electron.* **39**, 221 (2007)

Chapter 4

Waveguide Fabrication Techniques

In Chapter 3, the theoretical considerations relevant to various types of waveguides were discussed. In every case, waveguiding depended on the difference in the index of refraction between the waveguiding region and the surrounding media. A great many techniques have been devised for producing that required index difference. Each method has particular advantages and disadvantages, and no single method can be said to be clearly superior. The choice of a specific technique of waveguide fabrication depends on the desired application, and on the facilities available. In this Chapter, various methods of waveguide fabrication are reviewed, and their inherent features are discussed.

4.1 Deposited Thin Films

One of the earliest-used, and most effective, methods of waveguide fabrication is the deposition of thin films of dielectric material. In recent years glass and polymer waveguides have greatly increased in importance because they are inexpensive compared to those made of semiconductors or lithium niobate. Millions of inexpensive waveguide devices are needed to fully implement the international-communications network, sometimes called the “information superhighway”. In this section, the word deposition is broadly defined to include methods of liquid-source deposition, such as spinning or dipping, as well as vacuum-vapor deposition and sputtering. Because of the rapidly expanding use of polymer films in waveguide devices an entire chapter has been devoted to them in this edition. The fabrication of polymer waveguides and their use in a variety of integrated optic applications is described in Chapter 5.

4.1.1 Sputtered Dielectric Films

Thermally stimulated vacuum-evaporation, which is the standard method of producing thin films for conventional applications, such as anti-reflection coatings, is seldom used for waveguide fabrication, since it produces films with relatively high loss

at visible wavelengths (10 dB/cm). This high loss is due to inclusion of contaminant atoms that act as absorption and scattering centers [1]. Instead, sputter deposition of molecules from a solid source is used. Sputtering is the process whereby atoms or molecules are removed from the surface of a source (target) material, in vacuum, by bombardment with ions having energies in excess of about 30 eV to about 2 keV. The atoms (or molecules) removed from the surface of the target are allowed to deposit on the surface of a substrate, forming a thin film layer. The sputtered thin film is slowly built up by an accumulation of individual particles arriving at the surface with significant kinetic energy. This process produces a very uniform layer because the deposited atoms are kinetically distributed over the surface. Contaminant atoms are mostly excluded because the process is carried out at a relatively low temperature compared to that required for vacuum-vapor-deposition, and the target material can be highly purified before use. As a result, good quality optical films can be produced, with losses on the order of 1 dB/cm [2].

One method of sputtering [3] a thin film layer is by plasma discharge, as diagrammed in Fig. 4.1. The target and substrate are placed in a vacuum system and a gas is let in at a pressure of $(2\text{--}20) \times 10^{-3}$ torr. A high voltage bias is applied between anode and cathode so that a plasma discharge is established. Ions generated in the plasma discharge are accelerated toward the cathode. They strike the target, thus transferring their momentum to atoms near the surface of the target material, which are thereby sputtered off, and then deposited on the substrate.

In order for the sputtered atoms to adhere well to the substrate so as to produce a uniform, low-loss layer, the substrate must be thoroughly cleaned in an appropriate solvent or etchant prior to being placed in the vacuum system. Substrate cleaning is a basic step in any thin-film deposition process, and methods vary, depending on the substrate and film materials. A review of substrate cleaning procedures has been given by Zernike [4].

The position of the substrate in Fig. 4.1 is merely representative of one of the many different geometric configurations that can be used, since atoms sputtered

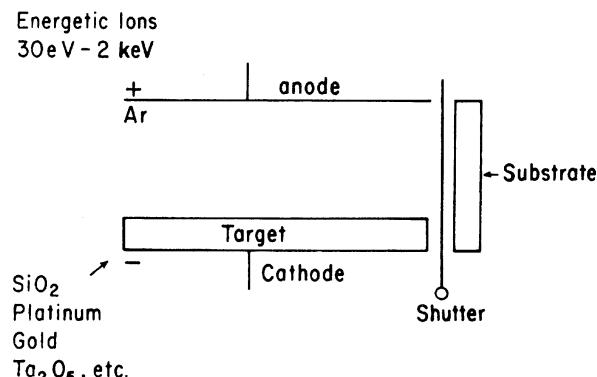


Fig. 4.1 Plasma discharge system for sputter deposition of thin films

from the target tend to deposit on every exposed surface within the vacuum system. For best uniformity of the deposited layer, the distance from substrate to target should be large compared to the substrate dimensions. However, increased substrate-target separation reduces the deposition rate, so that a longer time is required to produce a layer of a given thickness. Often the substrate is attached to the anode because it provides a parallel geometry that is conducive to layer uniformity. However, electrical isolation must be provided if a separate substrate bias is desired, and substrate heating due to incident electrons may be significant. The purpose of the shutter, shown in Fig. 4.1, is to shield the substrate from deposition during the first minute or so after the plasma discharge is activated, since many adsorbed contaminant atoms are released during that period.

Usually one of the inert gases, argon, neon, or krypton, is used in the plasma discharge chamber to avoid contamination of the deposited layer with active atoms. However, in some cases it is advantageous to employ reactive sputtering, in which atoms sputtered from the target react with the bombarding ions to form an oxide or nitride deposited film. For example, a silicon target can be used in the presence of ammonia gas to form silicon nitride according to the reaction



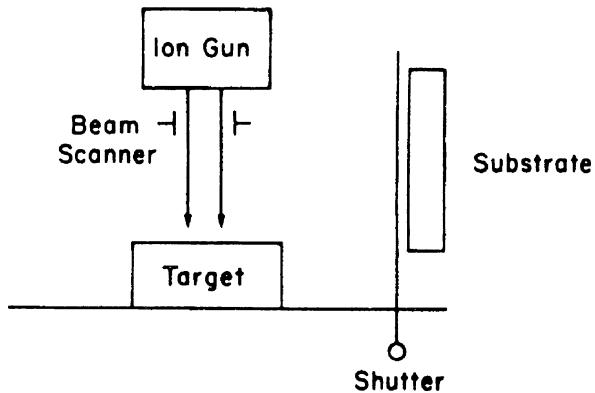
The bias voltage applied between anode and cathode may be either dc, or rf with a dc component. rf sputtering usually produces better quality films and is not subject to charge build-up problems in low conductivity targets but, of course, a more complicated power supply is needed. Detailed descriptions of both rf and dc sputtering are available elsewhere [5, 6].

Plasma discharge sputtering is also useful for depositing metal films, as well as dielectric layers. Thus, metallic electrical contacts and field plates can be deposited on top of, or adjacent to, dielectric waveguides, as required.

As an alternative to generating the bombarding ions in a plasma discharge, one can use a collimated beam of ions produced by a localized source, or ion gun, as shown in Fig. 4.2. The advantages of ion beam sputtering, as compared to the plasma discharged method, are that deposition can be done in a high vacuum $< 10^{-6}$ torr, and that the focused ion beam strikes only the target. No contaminant atoms are sputtered off of the walls of the chamber. Since the focused ion beam is often smaller than the target, the beam is usually electrically scanned in a raster pattern over the target to insure uniform sputtering.

Sputter deposition, whether plasma discharge or ion beam can be used to fabricate a wide variety of dielectric waveguides. To cite just a few examples. Zernike has produced Corning 7059 glass waveguides on both glass and KDP substrates [4], while Stadler and Gopinath have made magneto-optical films [7]. Choi et al. have sputtered erbium-doped SiO_2 to make waveguide optical amplifiers [8], while Kitigawa et al. have used sputtered neodymium-doped silica waveguides to make lasers [9]. Sputtered waveguides have also been fabricated from tantalum oxide [10] and yttrium oxide [11].

Fig. 4.2 Ion beam sputter deposition system



4.1.2 Deposition from Solutions

Many materials can be used to form waveguides by applying a solution which dries to form a dielectric film and by spinning the substrate on a photoresist spinner to spread the layer evenly, by dipping the substrate into the solution and slowly withdrawing it, or by casting or injection-molding techniques. Table 4.1 lists some common waveguide materials that can be applied in this fashion, along with the appropriate solvent in each case. The advantage of these waveguides is that they are inexpensive and can be applied without any sophisticated equipment. However, material purity is often low compared to that obtainable in sputtered films, and uniformity is relatively poor. Nevertheless, some surprisingly good results have been obtained. Neyer et al. [18] have used injection molding of polymethylmethacrylate (PMMA) filled with ethyleneglycoldimethacrylate (EGDMA) to form waveguides with 0.3 dB/cm loss at 1300 nm wavelength.

Polyimide channel waveguides fabricated by direct laser writing have been used as the optical interconnect medium between 50 GHz modulated sources and high-speed photodetectors [19]. Such interconnects could potentially be *utilized* in the implementation of sophisticated water-scale systems such as phased array radars.

Table 4.1 Waveguide materials applied by spinning and dipping

Material	Solvent
Photoresist [12]	Acetone
Epoxy [13]	Proprietary Compounds
Polymethylmethacrylate [14]	Chloroform, Toluene
Polyurethane [15]	Xylene
Polyimide [16]	
Spun-on glass (SOG) [17]	

4.1.3 Organosilicon Films

Tien et al. [20] have produced thin film waveguides by rf discharge polymerization of organic chemical monomers. Films prepared from vinyltrime-thylsilane (VTMS) and hexamethyldisiloxane (HMDS) monomers were deposited on glass substrates. The VTMS films had an index of 1.531 at 6328 Å, about 1% larger than the index of the substrate of ordinary glass (1.512). The corresponding index of HMDS films was 1.488; thus, a substrate of Corning 744 Pyrex glass, which has an index of 1.4704, was used. Deposition was accomplished by introducing monomers and argon through separate leak valves into a vacuum system which had been evacuated to a pressure of less than 2×10^{-6} torr. A discharge was then initiated by applying a 200 W, 13.56 MHz rf voltage between an anode and a cathode on which the substrate rested. VTMS films grew at a rate of about 2000 Å/min, while the rate for HMDS films was 1000 Å/min, for metered pressures of 0.3 torr (monomer) and 0.1 torr (Ar). The resulting polymer films were smooth, pinhole free, and transparent from 4000 Å to 7500 Å. Optical losses were exceptionally low (< 0.004 dB/cm). In addition, it was found that the index of the films could be varied either by mixing of the two monomers prior to deposition, or by chemical treatment after deposition.

4.2 Substitutional Dopant Atoms

While thin-film deposition techniques have proved very effective for producing waveguides in glasses and other amorphous materials, they generally cannot be used to produce layers of crystalline materials, such as semiconductors, or ferroelectrics like LiNbO_3 and LiTaO_3 . This is true because the elements of these compound materials usually do not deposit congruently; the relative concentrations of different elements are not maintained as atoms are transferred from source to substrate. Even when conditions can be established so that congruent deposition occurs, the grown layer will generally not be single crystal and epitaxial (of the same crystal structure as the substrate on which it is grown). To avoid this problem, many different approaches have been developed to produce waveguides in, or on, crystalline materials without seriously disrupting the lattice structure. A number of these methods involve the introduction of dopant atoms which substitutionally replace some of the lattice atoms, bringing about an increase in the index of refraction.

4.2.1 Diffused Dopants

Since waveguides for visible and near-infrared wavelength usually have thicknesses of only a few micrometers, diffusion of dopant atoms from the surface of the substrate is a viable approach to waveguide fabrication. Standard diffusion techniques are used [21], in which the substrate is placed in a furnace, at typically 700° to 1000°C, in the presence of a source of dopant atoms. This source may be a flowing

gas, liquid, or solid surface film. For example, diffusion of atoms from titanium, tantalum, nickel or zinc metal surface layers is often used to produce waveguides in LiNbO₃ and LiTaO₃ [22–25]. Metal diffused waveguides in LiNbO₃ and LiTaO₃ can be produced with losses that are smaller than 1 dB/cm.

The most commonly used method for producing waveguides in LiNbO₃ is to deposit a layer of titanium on the surface by rf sputtering or e-beam evaporation. Then the substrate is placed in a diffusion furnace at about 1000°C and heated in a flowing gas atmosphere containing oxygen. For example, argon which has been passed through an H₂O bubbler is often used. As the Ti coated LiNbO₃ substrate is moved into the furnace, oxidation of the deposited Ti begins when its temperature reaches about 500°C. By the time the substrate has reached the desired diffusion temperature, set somewhere in the range of 1000–1100°C, oxidation is complete and the Ti metal has been replaced by a layer of polycrystalline TiO₂. This TiO₂ subsequently reacts with the LiNbO₃ to form a complex compound at the surface which acts as the diffusion source of Ti atoms. The extent of Ti diffusion depends on the diffusion coefficient and on time and temperature. It follows the usual Fick's laws for diffusion from a limit source of atoms yielding a Gaussian distribution of diffused Ti atoms as a function of depth from the surface. The advantage of this multi-staged diffusion process, as compared to diffusion directly from a metallic Ti layer, is that the out-diffusion of Li atoms from the LiNbO₃ is prevented. Li out-diffusion would otherwise result in changes in the index of refraction to a much greater depth than that desired for single mode waveguides. With the oxidation process, the change in index occurs only over the depth of the diffused Ti profile, which is accurately controllable.

Metal diffusion can also be used to produce waveguides in semiconductors [26]. Generally, a p-type dopant is diffused into a n-type substrate, or vice versa, so that a p-n junction is formed, thus providing electrical isolation as well as optical waveguiding. For example, Martin and Hall [27] have used diffusion of Cd and Se into ZnS and Cd into ZnSe to make both planar and channel waveguides. Losses in 10 μm wide, 3 μm deep channel waveguides formed by diffusion of Cd into ZnSe were measured to be less than 3 dB/cm.

Diffusion of substitutional dopant atoms to produce waveguides is not limited to semiconductors; it can also be done in glass and polymers. For example, Zolatov et al. [28] have studied waveguides fabricated by diffusing silver into glass. The observed change in index of refraction resulting from the presence of the diffused Ag atoms was $\Delta n \cong 0.073$ at wavelengths from 6328 Å to 5461 Å. Because diffusion must be performed below the relatively low temperatures at which glasses melt, it is a slow process. Hence, the alternative methods of ion exchange and migration are often used to introduce the required dopant atoms.

Many polymers have excellent dopant-diffusion capabilities, as described by Booth [29]. Solvent bath dopants can be selectively diffused into guide regions to increase the index, or low molecular weight dopants or monomers can be thermally out-diffused to create guide regions.

4.2.2 Ion Exchange and Migration

A typical ion migration and exchange process is diagrammed in Fig. 4.3. The substrate material is a sodium-doped glass. When an electric field is applied as shown, and the glass substrate is heated to about 300°C, Na^+ ions migrate toward the cathode. The surface of the glass is submerged in molten thallium nitrate. Some of the Na^+ ions are exchanged for Tl^+ , resulting in the formation of a higher index layer at the surface.

Izawa and Nakagome [30] have used an ion exchange and migration method to form a buried waveguide slightly beneath the surface. Their setup is similar to that of Fig. 4.3 except that a borosilicate glass substrate is used, at a temperature of 530°C. A two stage process is employed. In the first stage, the salts surrounding the substrate (anode) are thallium nitrate, sodium nitrate and potassium nitrate, so that Na^+ and K^+ ions in the glass are exchanged with Tl^+ ions, forming a higher-index layer at the surface. In the second stage, only sodium nitrate and potassium nitrate are present at the anode. Hence, the Na^+ and K^+ ions diffuse back into the surface, while the Tl^+ ions move deeper into the substrate. The resulting structure has a higher-index, Tl^+ doped layer located at some depth below the surface, covered on top and bottom by the normal borosilicate glass. Thus, a *buried* waveguide has been formed. Ion exchange can also be used to produce waveguides in LiNbO_3 . In that case, protons are exchanged for lithium ions to produce the desired increase in index of refraction. The source of protons, benzoic acid, is heated above its melting point (122°C) to a temperature of about 235°C. The masked lithium niobate substrate is then placed in the melt for a period of one hour or less. This treatment is sufficient to produce a step-like index change of (extraordinary) $\Delta n_e = 0.12$ and (ordinary) $\Delta n_o = -0.04$ [31]. Design procedures and software tools for designing ion exchange waveguides have been published [32]. Ion exchange waveguides have been used to make a variety of devices, including an add/drop filter [33], an arrayed-waveguide grating multi/demultiplexer [34] and a periodically segmented waveguide [35].

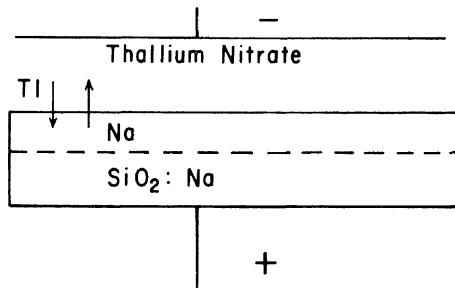


Fig. 4.3 Ion migration system for waveguide fabrication

4.2.3 Ion Implantation

Substitutional dopant atoms can also be introduced by ion implantation. In the ion implantation doping process, ions of the desired dopant are generated, then accelerated through, typically, 20–300 keV before impinging on the substrate. Of course, the process must be carried out in a vacuum. A basic system for implantation doping is shown in Fig. 4.4. Detailed descriptions of ion implantation doping techniques are available elsewhere [36]; hence, they will not be repeated here. However, it should be noted that the basic elements of the implantation system are an ion source, accelerating electrodes, electrostatic or magnetic deflection elements combined with a slit to form an ion (mass) separator, and a beam deflector to scan the collimated ion beam in a raster pattern over the substrate. The depth of penetration of the implanted ions depends on their mass, energy, and on the substrate material and its orientation. Tabulated data are available for most ion/substrate combinations that are of interest [37]. Knowing the penetration characteristics and the implanted ion dose per cm^2 , which can be determined very accurately by measuring beam current density and implant time, one can calculate the profile of implanted ion concentration with depth.

In most cases, ion implantation must be followed by annealing at elevated temperatures to remove implantation-caused lattice damage and to allow the implanted dopant atoms to move into substitutional sites in the lattice. After annealing, waveguides made by substitutional dopant implantation have much the same characteristics as diffused waveguides. However, the implantation process allows greater control of the dopant concentration profile, since ion energy and dose can be varied to produce either flat or other desired distributions. Diffusion always yields either a Gaussian or a complementary error function distribution of dopant atoms, depending on the type of source used.

A side from the substitutional dopant effect, ion implantation can also be used to produce waveguides in which the increase in index of refraction is the result of lattice disorder produced by non-substitutional implanted ions. An example of this type of waveguide has been produced by Standley et al. [38]. They found that implantation of fused quartz substrates with a variety of ions ranging from helium to bismuth produced a layer of increased index. Ion energy was in the range of 32–200 keV. The best results were obtained using lithium atoms, for which the empirical relation between ion dose and index of refraction at 6328 Å was given by

$$n = n_0 + 2.1 \times 10^{-21} C, \quad (4.2)$$

where n_0 is the preimplantation index and C is the ion concentration per cm^3 . Optical losses less than 0.2 dB/cm were achieved after annealing for 1 hr at 300°C.

A third type of optical waveguide can be produced by ion implantation in certain semiconductors that are subject to defect center trapping. Proton bombardment can be used to generate lattice damage, which results in the formation of compensating centers, to produce a region of very low carrier concentration in a substrate with relatively large carrier concentration. The index of refraction is lightly larger in the

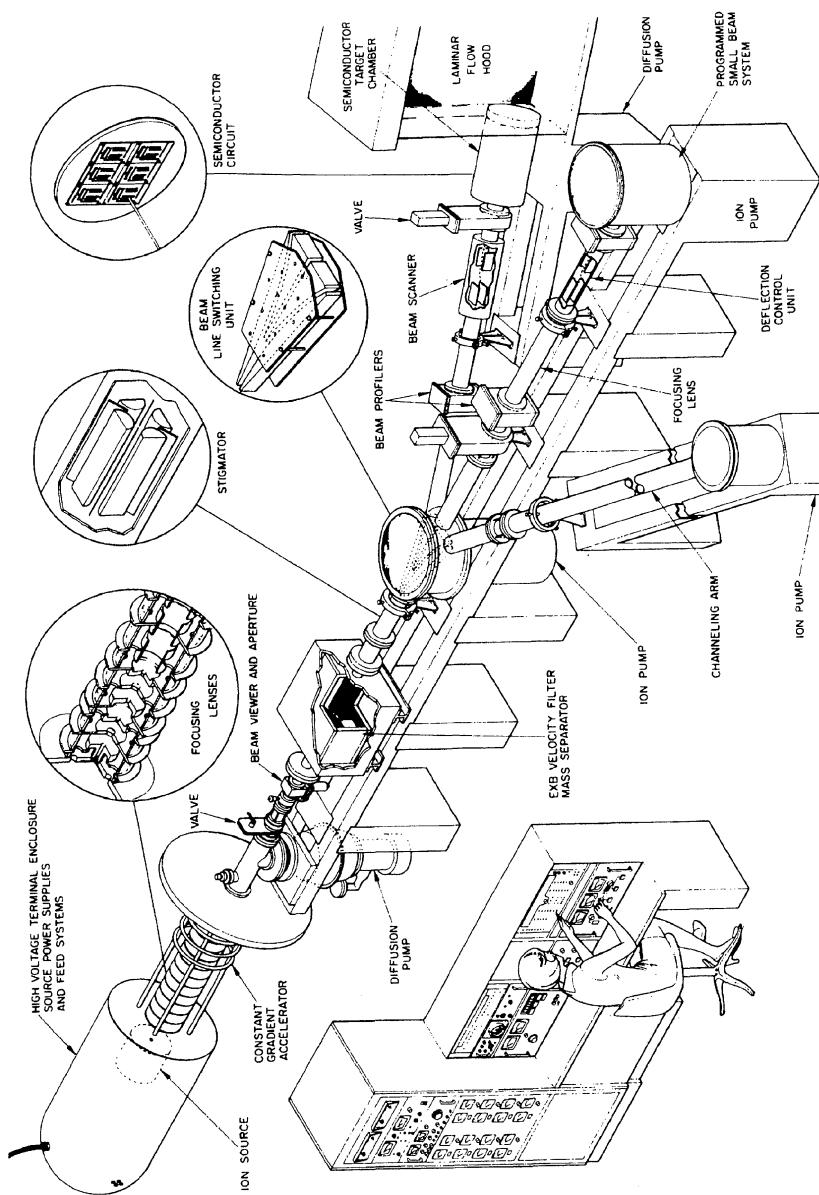


Fig. 4.4 Drawing of a 300-kV ion-implantation system with three arms for separate applications of the ion beam (Diagram courtesy of Hughes Research Laboratories, Malibu, CA)

low carrier concentration region, because free carriers normally reduce the index. The implanted protons do not by themselves cause a significant increase in index; the index difference results from carrier trapping which lowers the carrier concentration. This type of waveguide, which was first demonstrated in GaAs [39, 40], is discussed in detail in Section 4.3, along with other types of carrier-concentration-reduction waveguides.

4.3 Carrier-Concentration-Reduction Waveguides

In a semiconductor, any free carriers that are present contribute negatively to the index of refraction; i.e., they reduce the index below the value that it would have in a completely carrier-depleted sample of the material. Thus, if the carriers are somehow removed from a region, that region will have a larger index than the surrounding media and could function as a waveguide.

4.3.1 Basic Properties of Carrier-Concentration-Reduction Waveguides

A quantitative expression for the reduction in index of refraction produced by a given concentration per cm^3 of free carriers, N , can be developed by analogy with the change in index produced in a dielectric by a plasma of charged particles [40]. In that case

$$\Delta \left(\frac{\varepsilon}{\varepsilon_0} \right) = 2n\Delta n = -\frac{\omega_p^2}{\omega^2} = -\frac{Ne^2}{\varepsilon_0 m^* \omega^2}, \quad (4.3)$$

where ω_p is the plasma frequency, m^* is the effective mass of the carriers, and the other parameters are as defined previously. If we define n_0 as the index of refraction of the semiconductor when free carriers are absent, the index with N carriers per cm^3 present is

$$n = n_0 - \frac{Ne^2}{2n\varepsilon_0 m^* \omega^2}. \quad (4.4)$$

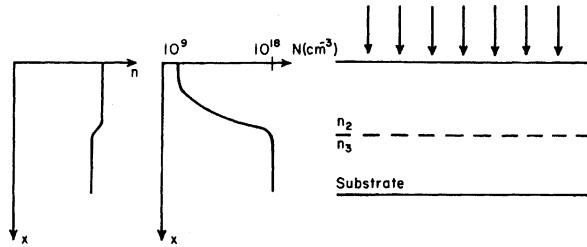
Thus, the difference in index between a waveguiding layer with N_2 carriers per cm^3 and a confining layer with N_3 carriers per cm^3 is

$$\Delta n = n_2 - n_3 = \frac{(N_3 - N_2)e^2}{2n_2\varepsilon_0 m^* \omega^2}, \quad (4.5)$$

where $N_3 > N_2$, and where it has been assumed that $n_2 \approx n_3$.

Some interesting features of carrier-concentration-reduction waveguides are evident in (4.5). First, it can be seen that once N_2 has been reduced by a factor of about

Fig. 4.5 Diagram of proton bombardment method of waveguide fabrication, showing the resultant carrier concentration (N) and index of refraction (n) profiles with depth



10 below the value of N_3 , further reduction does not substantially change the index difference Δn . The significance of this fact is that proton bombarded waveguides of the carrier-concentration-reduction type have a much more abrupt change in index at the interface between the waveguide and the confining layer than might be expected from the distribution profile of bombardment-created damage. This effect is illustrated in Fig. 4.5.

A second interesting feature of carrier-concentration-reduction waveguides can be seen by considering an asymmetric waveguide like that of Fig. 3.2, with $n_1 \ll n_2, n_3$. In that case, the carrier concentration difference required to permit waveguiding of the lowest order ($m = 0$) mode is found by combining (4.5) and (3.31), which yields

$$N_3 - N_2 \geq \frac{\varepsilon_0 m^* \omega^2 \lambda_0^2}{16e^2 t_g^2}. \quad (4.6)$$

Since $\omega = 2\pi v$ and $v\lambda_0 = c$, (4.4) can be written

$$N_3 - N_2 \geq \frac{\pi^2 c^2 m^* \varepsilon_0}{4e^2 t_g^2}. \quad (4.7)$$

Notice that the cutoff condition of (4.7) is independent of wavelength! Thus, in the case of a carrier-concentration-reduction type waveguide, no greater difference in concentration is required to guide light at, say, 10 μm wavelength than that required for 1 μm light. Of course, the shape of the mode, particularly the size of the evanescent tail extending into Region 3, would be different, which would result in different losses at the two wavelengths. The reason for the unusual lack of dependence of the cutoff condition on wavelength is that both the required Δn and the Δn produced by a given value of $N_3 - N_2$ are proportional to λ_0^2 , as can be seen from (4.6) and (3.31).

4.3.2 Carrier Removal by Proton Bombardment

Carrier-concentration-reduction waveguides have been produced by proton bombardment in both GaAs and GaP [40–42]. In general, the optical losses are greater than 200 dB/cm after bombardment with a proton dose that is greater than 10^{14} cm^{-2} . However, after annealing at temperatures below 500°C, losses can be reduced below 3 dB/cm [39]. Typical anneal curves for samples of GaAs implanted with 300 keV protons are shown in Fig. 4.6. After annealing to remove excess optical absorption associated with the damage centers, the remaining loss is mostly due to free carrier absorption in the tail of the mode that extends into the substrate (see Fig. 2.5 for a representation of modes in proton bombarded waveguide). Substitution of typical values of m^* for GaAs and GaP, respectively, into (4.7) indicates that, for $t_g = 3 \mu\text{m}$, substrate carrier concentrations of $N_3 = 6 \times 10^{17} \text{ cm}^{-3}$ and $N_3 = 1.5 \times 10^{18} \text{ cm}^{-3}$ are required to guide even the lowest-order mode. Hence, free carrier absorption in the mode substrate tail can be a serious problem. Increasing the ratio t_g/λ_0 reduces the relative size of the mode tail, mitigating the effect of related free carrier absorption. By careful control of ion dose and/or anneal conditions it is possible to produce proton bombarded waveguides with losses less than 2 dB/cm for use at wavelengths ranging from 1.06 μm to 10.6 μm [40].

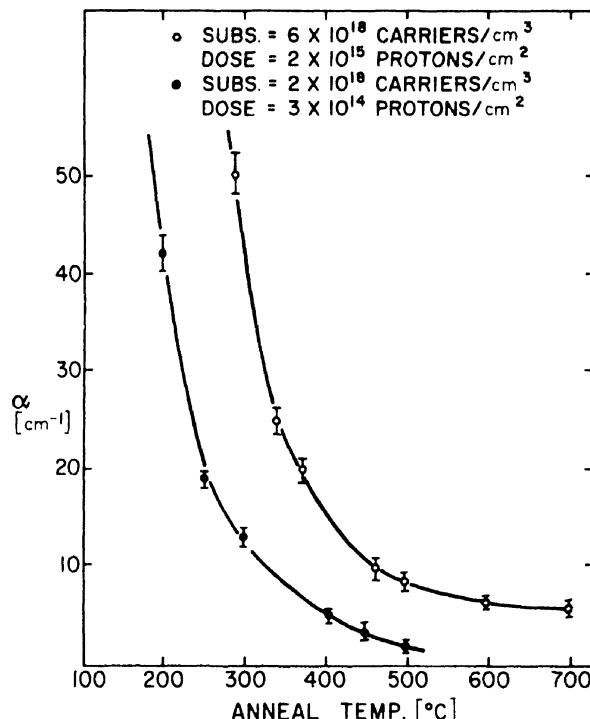


Fig. 4.6 Loss in proton bombarded waveguides. (Data are given in terms of the exponential loss coefficient α , defined by $I(z) = I_0 \exp(-\alpha z)$, where I_0 is the optical intensity at the starting point, $z = 0$)

In addition to GaAs and GaP, ZnTe [43] and ZnSe [44] are materials in which proton-bombardment generated carrier compensation can be used to produce waveguides. The method is effective in both n- and p-type substrates, suggesting that deep level traps are involved rather than shallow acceptors or donors. Proton bombarded waveguides are useful over the normal transparency range of the substrate material. Thus, for example, waveguides for 6328 Å wavelength can be made in GaP, but not in GaAs. Waveguides for 1.06 μm, 1.15 μm and 10.6 μm have been made in both GaAs and GaP.

4.4 Epitaxial Growth

In a monolithic optical integrated circuit formed on a semiconductor substrate, epitaxial growth is the most versatile method of fabricating wave guides. This is true because the chemical composition of the epitaxially grown layer can be varied to adjust both the index of refraction and the wavelength range of transparency of the waveguide [45].

4.4.1 Basic Properties of Epitaxially Grown Waveguides

In monolithic semiconductor OIC's, there exists a basic problem of wavelength incompatibility that must be overcome. A semiconductor has a characteristic optical emission wavelength, corresponding to its bandgap energy, which is approximately the same as its absorption-edge wavelength. Thus, if a light emitting diode or laser is fabricated in a semiconductor substrate, it will emit light that is strongly absorbed in a waveguide formed in the same substrate material. Also, this light will not be efficiently detected by a detector formed in that substrate because its wavelength will correspond to the *tail* of the absorption-edge. In order to produce an operable OIC, the effective bandgap energies for absorption and emission must be altered in the various elements of the circuit so that

$$E_g \text{ waveguide} > E_g \text{ emitter} > E_g \text{ detector.} \quad (4.8)$$

Epitaxial growth of a ternary (or quaternary) material offers a convenient means of producing this required change of bandgap energy, as well as a change in index of refraction. In general, change in bandgap on the order of tenths of an eV can be produced by changing the atomic composition of the grown layer by as little as 10%. Corresponding changes in index of refraction occur attendantly. With careful design, one can fabricate waveguides that have low loss at the emitted wavelength, along with efficient detectors. The epitaxial growth technique can best be illustrated by describing its application to $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$, which is the most commonly used material for monolithic OIC fabrication.

4.4.2 $Ga_{(1-x)}Al_xAs$ Epitaxially Grown Waveguides

The growth of $Ga_{(1-x)}Al_xAs$ epitaxial layers on a GaAs Substrate for OIC fabrication is usually carried out by liquid phase epitaxy (LPE), in a tube furnace, at temperatures in the range from 700 to 900°C. A more-or-less standard approach has been developed in which a number of liquid melts, of relatively small volume ($\sim 0.1 \text{ cm}^3$), are contained in a movable graphite *slidebar* assembly that allows the substrate to be transported sequentially from one melt to the next so as to grow a multi-layer structure. Details of this approach have been reviewed by Garmire [46].

Epitaxial growth of multi-layered $Ga_{(1-x)}Al_xAs$ structures with different Al concentration in each layer, as shown in Fig. 4.7, can be used to create waveguides with transparency for light of wavelengths longer than the bandgap (absorption-edge) wavelength, which varies depending on Al concentration, as shown in Fig. 4.8.

The curves in Fig. 4.8 were generated by shifting the experimentally measured absorption edge of GaAs [47, 48] by an amount corresponding to a change in bandgap calculated from [49]

$$E_g(x) = 1.439 + 1.042x + 0.468x^2, \quad (4.9)$$

where x is the atomic fraction of Al in the $Ga_{(1-x)}Al_xAs$. Al concentrations $> 35\%$ are not often used because above that level the bandgap is indirect, resulting in deleterious effects which are discussed in Chapter 10. Thus, the Al concentration x can be selected so that the waveguide is transparent at the desired wavelength.

Optical losses as low as 1 dB/cm are obtainable. Once the Al concentration in the waveguiding layer has been selected, the next step in the waveguide design is to select the Al concentration y in the confining layer.

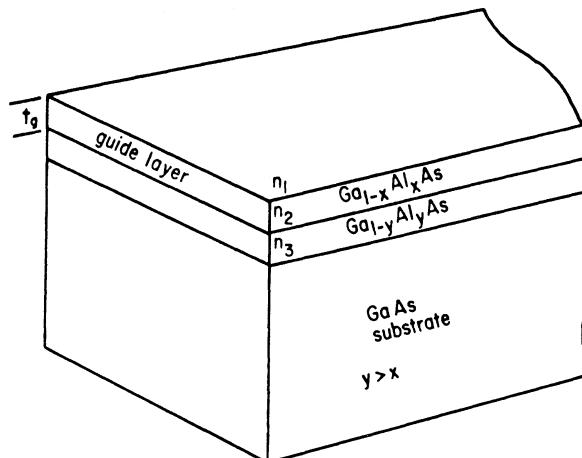


Fig. 4.7 $Ga_{(1-x)}Al_xAs$ planar waveguide

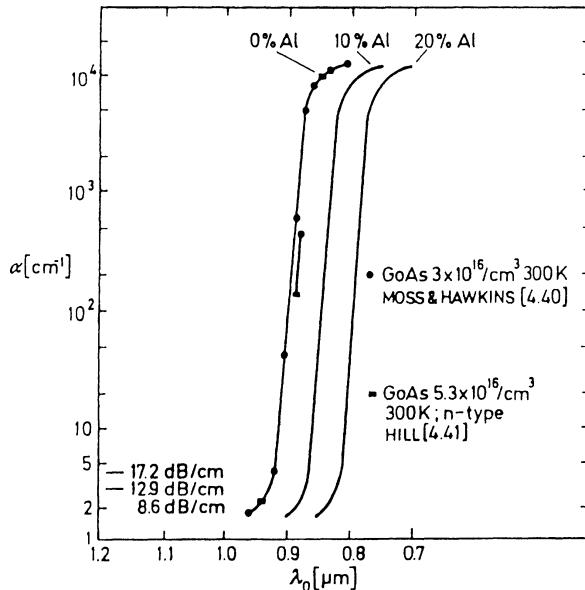


Fig. 4.8 Interband absorption as a function of wavelength and Al concentration

Increasing the Al concentration in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ causes the index of refraction to decrease. The dependence is shown in Fig. 4.9, which is a plot of the empirically determined Sellmeier equation [50, 51]

$$n^2 = A(x) + \frac{B}{\lambda_0^2 - C(x)} - D(x)\lambda_0^2, \quad (4.10)$$

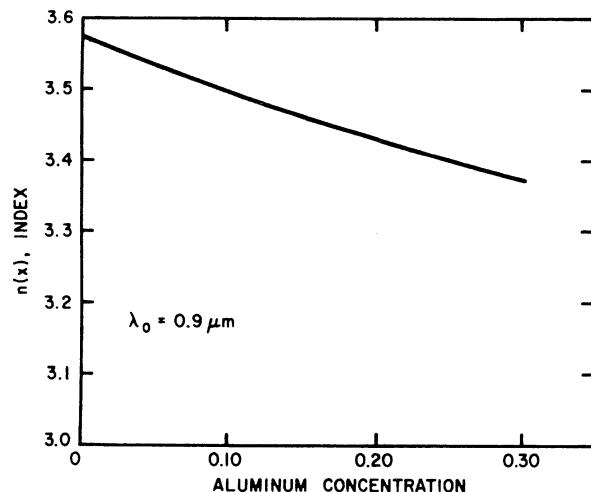


Fig. 4.9 Index of refraction of $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ as a function of Al concentration

Table 4.2 Sellmeier equation coefficients (for λ_0 in μm)

Material	A	B	C	D
GaAs	10.906	0.97501	0.27969	0.002467
$\text{Ga}_{(1-x)}\text{Al}_x\text{As}$	$10.906 - 2.92x$	0.97501	$(0.52886 - 0.735x)^2$ $x \leq 0.36$ $(0.30386 - 0.105x)^2$ $x \geq 0.36$	$(0.002467)(1.41x + 1)$

where x is the atomic fraction of Al atoms in the $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$, and where A, B, C, D are functions of x as given in Table 4.2. Because of the slight nonlinearity of the n vs. x curve, the index difference between two layers with unequal Al concentrations depends not only on the difference in concentration, but also in its absolute level. This effect is illustrated in Fig. 4.10 for two different Al concentrations in the waveguiding layer, 0 and 20%, respectively.

For the case of an asymmetric $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ waveguide, as shown in Fig. 4.7, the cutoff condition (3.31) can be used in conjunction with (4.10) to calculate what Al concentration is required to produce waveguiding of a given mode. Figure 4.11 shows the results of such a calculation for the case of waveguiding of the two lowest order modes of $0.9\ \mu\text{m}$ wavelength light in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ waveguides of various

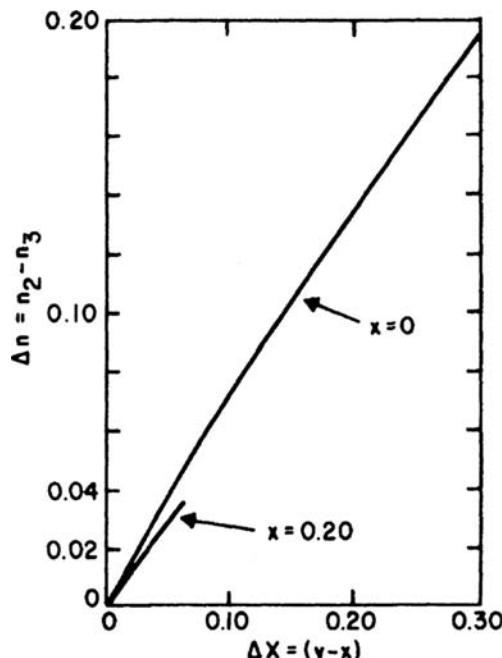


Fig. 4.10 Refractive index difference as a function of Al concentration difference between two layers of $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ for two values of absolute Al concentration

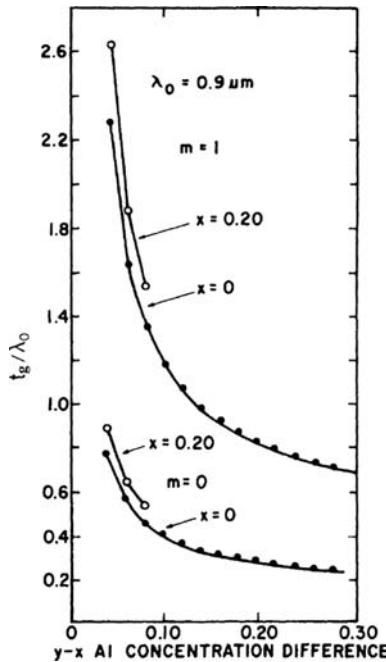


Fig.4.11 Guide thickness to wavelength ratio as a function of Al concentration between layer and substrate

thicknesses and Al concentrations [52]. Note the sharp increase in Al concentration difference ($y - x$) required for thickness/wavelength ratios less than about 0.8, particularly in the case of the $m = 1$ mode. However, for $t_g/\lambda_0 \gtrsim 1.0$, a concentration difference of only 10% is sufficient to waveguide both the $m = 0$ and the $m = 1$ mode.

$\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ multilayer waveguides, produced by epitaxial growth, have been used in a variety of optical devices, of both discrete and integrated form. Alferov et al. [53] have used this type of waveguide to make one of the earliest confined field lasers, a topic that will be discussed in greater detail in Chapter 12. $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ epitaxial waveguides have been used to make modulators and beam deflectors for IR light of $10.6 \mu\text{m}$ wavelength [54], and they also have been utilized in optical integrated circuits for $1 \mu\text{m}$ wavelength [55, 56].

To conclude this section on $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ waveguides grown by liquid phase epitaxy, it should be noted that alternatives to the limited-melt, slidebar growth method are available. For example, Crawford and Groves [57] have grown $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ layers by vapor phase epitaxy (VPE), and Kamath [58] has used a vertical dipping *infinite melt* technique of LPE. The major advantage of both of these methods is that they permit the use of large area substrates. The most promising large-area-substrate epitaxial growth method is molecular beam epitaxy (MBE) which is discussed in Section 4.4.4.

4.4.3 Epitaxial Waveguides in Other III-V and II-VI and IV Materials

Most of the work done on epitaxially grown OIC's has been done in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ because of the extraordinary close match of the lattice constants of GaAs and AlAs, 5.64 Å and 5.66 Å, respectively [59]. However, other III-V and II-VI materials are of interest. Waveguides have been grown in $\text{Ga}_{(1-x)}\text{In}_x\text{As}$ [60] and in $\text{CdS}_x\text{Se}_{(1-x)}$ [61]. $\text{Ga}_{(1-x)}\text{In}_x\text{As}$ is particularly interesting because the addition of In to GaAs shifts the bandgap, and hence the absorption-edge, to longer wavelengths. This suggests the possibility of $\text{Ga}_{(1-x)}\text{In}_x\text{As}$ OIC's in which the operating wavelengths of emitters and detectors could be made longer by the addition of indium, while more pure GaAs could be used for waveguides to minimize absorption. Such circuits would be particularly useful in conjunction with fiber-optic interconnections, since the minimum absorption of glass fibers occurs for about 1.2 μm wavelength [59]. Unfortunately, the lattice constants of GaAs and InAs do not match very well, being 6.06 Å for InAs as opposed to 5.64 Å for GaAs [46, p.257]. This mismatch results in serious lattice strain at the interfaces between layers of different In concentration, thereby increasing optical scattering and introducing phase distortion into the waveguided modes. This problem can be mitigated by using layers of the quaternary compound $\text{Ga}_{(1-x)}\text{In}_x\text{As}_y\text{P}_{(1-y)}$, which provides an additional degree of freedom in simultaneously optimizing both wavelength and lattice matching. The measured bandgaps of lattice matched $\text{In}_{(1-x)}\text{Ga}_x\text{As}_{(1-y)}\text{P}_y$ alloys as a function of alloy composition are shown in Fig. 4.12 [62]. The dotted curves give the bandgap as a function of phosphorous concentration, while the solid curve shows the concentrations of Ga(x) and P(y) required for the lattice matched

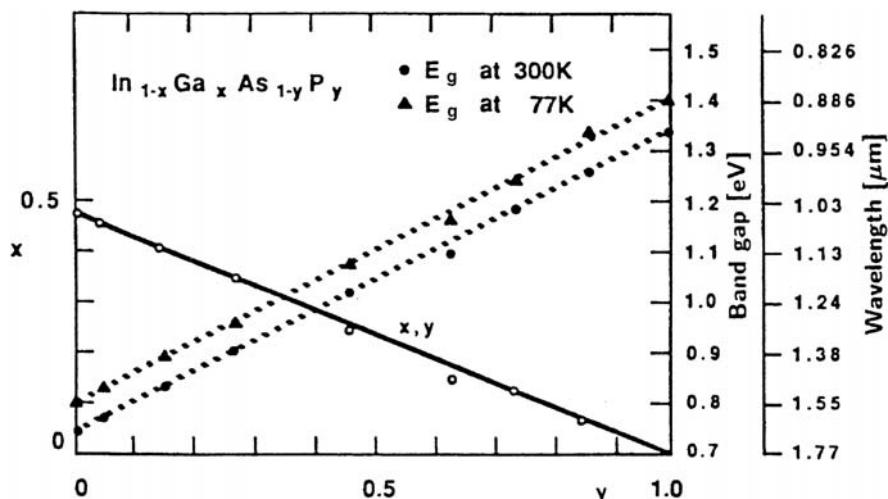


Fig.4.12 Measured band gaps of lattice matched $\text{In}_{(1-x)}\text{Ga}_x\text{As}_{(1-y)}\text{P}_y$ alloys as a function of alloy composition [62]

condition. To use these curves for light emitter design, one projects horizontally from the bandgap scale at the desired wavelength to the dotted curve corresponding to the desired operating temperature. The intersection determines the P concentration (y). (The corresponding As concentration is given by $1-y$.) The concentration of Ga(x) required for lattice matching is then read from the solid x,y curve at the point corresponding to the desired P concentration (y), using the values of x shown on the left vertical scale. The corresponding In concentration is given by $1-x$. Because of lattice matching and bandgap control obtainable, InGaAsP heterojunction lasers have become prevalent in many applications today, operating at 1300 and 1550 nm wavelengths. InGaAsP lasers have also been made to operate at 2550 nm [63]. The integration of such lasers with InGaAsP waveguides (and photodetectors) provides the basis for many optical integrated circuits. Recently epitaxial growth has been used to produce waveguides in SiGe [64].

4.4.4 Molecular Beam Epitaxy

In addition to the liquid phase epitaxy described in the preceding paragraphs, multilayered structures of semiconductors can also be grown by the technique of molecular beam epitaxy (MBE). In MBE, the constituent atoms are delivered to the substrate surface during growth by beams of accelerated molecules (or atoms). Since neither atoms nor molecules are charged particles they can only be accelerated by thermal excitation. A schematic diagram of an MBE growth chamber is shown in Fig. 4.13.

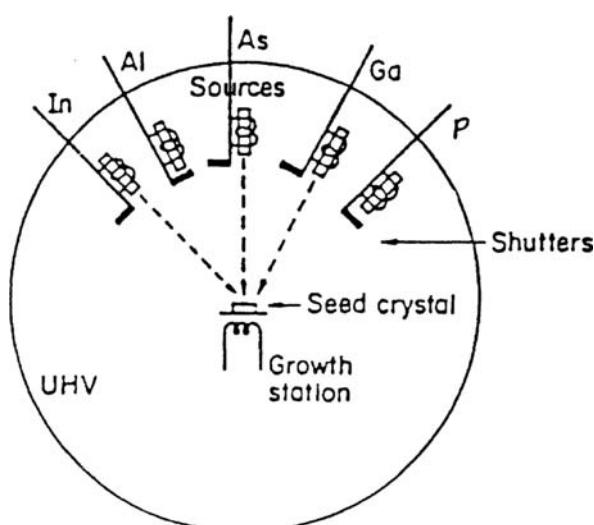


Fig. 4.13 A schematic of an MBE system

Sources for the growth of either GaAlAs or GaInAsP are shown. These sources consist of small, resistively heated crucibles that have an elongated shape with an opening at one end to produce a directed beam of thermally accelerated atoms. Each source has a remotely controlled shutter so that the beam can be turned on or off as desired at different points in the growth cycle. The entire assembly is enclosed in an ultra-high-vacuum chamber (UHV) so that the atoms are not scattered by atmospheric molecules. For a thorough discussion of the techniques of molecular beam epitaxy the reader is referred to the books by Chernov [65] and Ploog and Graf [66, 67].

The key advantage of MBE as compared to conventional LPE growth is that it provides superb control over purity, doping and thickness. To obtain the best control over purity, MBE systems generally consist of three separate chambers, a loading chamber, a growth chamber and an analysis chamber. This arrangement also increases the wafer throughput of the system because growth can be continuing on one substrate wafer while another is being loaded and prepared for growth. Remote control transfer mechanisms are used to move the wafer from one chamber to another [68]. In a carefully designed MBE system control of doping concentration and layer thickness with uniformity of $\pm 1\%$ over a 2-inch diameter wafer can be achieved [69]. Layers of thickness less than 100 Å can be controllably produced by MBE, permitting one to fabricate multilayered superlattices for use in quantum well devices [70]. This topic is discussed more thoroughly in Chapter 18.

4.4.5 Metal-Organic Chemical Vapor Deposition

Another technique of epitaxial growth that has been applied to opto-electronic device fabrication is metal-organic chemical vapor deposition (MOCVD), sometimes called metal-organic vapor phase epitaxy (MOVPE) [71–74]. In the case of MOCVD, the constituent atoms are delivered to the substrate as a gaseous flow within a growth reactor furnace. Metal-organic gases such as PH₂, AsH₂, TEGa, TEIn and THAl are used. The substrate temperature during growth is typically 750°C, a relatively low temperature compared to those used for LPE. The growth reactor can be a simple tube furnace such as those used for dopant diffusion. However, very complex (and expensive) control valves, filters and exhaust vents must be used to regulate the gaseous flow. These are required for the sake of safety because the Metal-organic gases are extremely toxic and the carrier gas (H₂) is explosive. Despite this inconvenience, MOCVD is growing in popularity as a method for fabricating opto-electronic devices because, like MBE, it offers precise control of dopant concentration and layer thickness. Quantum well structures can be grown by MOCVD. In addition, MOCVD can be used with large area substrate wafers. This latter property makes MOCVD very attractive as a method for large-volume production of lasers, LEDs and other opto-electronic devices.

4.5 Electro-Optic Waveguides

GaAs and $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ exhibit a strong electro-optic effect, in which the presence of an electric field produces a change in the index of refraction. Thus, if a metal field plate is formed over a GaAs substrate, so as to produce a Schottky barrier contact, as shown in Fig. 4.14, and a voltage is applied so as to reverse bias the Schottky barrier, the electric field in the depletion layer can cause a sufficient change in the index of refraction to produce a waveguiding layer, with index n_2 larger than that of the substrate, n_1 . Fortunately, it is very easy to form a Schottky barrier on GaAs or $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$. Almost any metal, except silver, will form a Schottky barrier rather than an ohmic contact, when deposited on n-type material without any heat treatment. When a reverse-bias voltage is applied to the Schottky barrier, a depletion layer forms, as in p-n junction. The index of refraction is increased in this layer by two different mechanisms.

First, the carrier depletion causes an increase as discussed in Sect. 4.3. Secondly, the presence of the electric field causes a further increase. For a detailed description of this linear electro-optic effect see, for example, Yariv [75]. The change in index of refraction for the particular orientation shown in Fig. 4.14, and for TE waves, is given by

$$\Delta n = n^3 r_{41} \frac{V}{2t_g}, \quad (4.11)$$

where n is the index of refraction with no field present, V is the applied voltage, t_g is the thickness of the depletion layer, and r_{41} is the element of the electro-optic tensor [75] appropriate to the chosen orientations of crystal and electric field. The electro-optic effect is non-isotropic, and other orientations will not necessarily exhibit the same change in index. For example, in the case of a crystal oriented as in Fig. 4.14 but for TM waves (with the E vector polarized in the x direction), the change in index is zero. Thus, both the orientations of the crystal substrate and the polarization of the waves to be guided must be carefully considered in the design of an electro-optic waveguide.

The thickness of the waveguide, i.e. the depletion layer thickness, depends on the substrate carrier concentration, as well as on the applied voltage. Assuming a

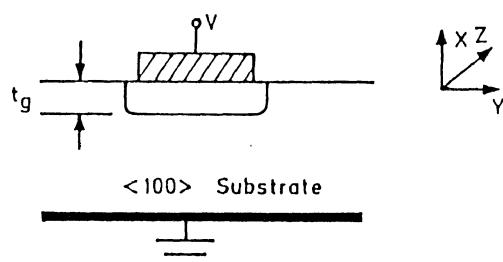


Fig. 4.14 Diagram of a basic electro-optic waveguide

reasonable substrate carrier concentration of 10^{16} cm^{-3} , application of $V = 100 \text{ V}$, i.e., the maximum allowed if avalanche breakdown is to be avoided, would produce a depletion layer thickness $t_g = 3.6 \mu \text{ m}$. This result is based on an abrupt junction calculation [76]. In that case, the resulting change in index, from (4.11), would be $\Delta n = 8.3 \times 10^{-4}$, where we have used $n = 3.4$ and $n^3 r_{41} = 6 \times 10^{-11} \text{ m/V}$ for GaAs [52]. From (3.31), it can be seen that the lowest order mode would be guided in this case, but just barely. ($\Delta n = 5.7 \times 10^{-4}$ is required.)

From the foregoing example, it is obvious that the constraints placed on waveguide design by the necessity of avoiding avalanche breakdown are serious. Unless one has available unusually lightly doped substrate material (carrier concentration $< 10^{16} \text{ cm}^{-3}$), only the lowest order mode can be guided. This problem can be alleviated by growing a relatively thick ($> 10 \mu\text{m}$) epitaxial layer of very lightly doped GaAs on the substrate before depositing the Schottky contact. Such layers can be grown with carrier concentration as low as 10^{14} cm^{-3} . In this case, care must be taken to use a low enough substrate carrier-concentration so that the epilayer does not form a carrier-concentration-reduction type waveguide.

The greatest advantage of the electro-optic waveguide, as compared to the others which have been described, is that it is electrically switchable. Thus, it can be used in switches and modulators, as is described in more detail in Chapter 8. Varying the applied voltage changes not only the index of refraction in the waveguiding region, but also the thickness of the waveguide, thus bringing the guide either above or below cutoff for a particular mode, as desired.

The field plate Schottky contact can be in the form of a narrow stripe, either straight or curved. In that case, a rectangular channel waveguide is formed underneath the stripe when the voltage is increased above the cutoff threshold level. Some additional methods of fabricating channel, or stripe, waveguides are described in the next section.

4.6 Oxidation

In the past fifteen years there has been growing interest in the fabrication of optical waveguides on silicon substrates as part of the effort to couple sophisticated silicon electronic integrated circuits to optical interconnects. It is well known that the native oxide of silicon, SiO_2 , can be grown very easily by heating the silicon to about 1000°C in the presence of either oxygen or water vapor. The resulting oxide has good uniformity, clings tenaciously to the silicon surface and can be grown to a thickness of several micrometers without cracking. For a thorough description of the silicon oxidation processes that are commonly used in the semiconductor industry see, for example, Campbell [77]. The thermal oxidation of silicon has been applied to the fabrication of optical waveguides on silicon substrates [78–81]. These waveguides exhibited good optical properties with relatively low loss.

The thermal oxidation processes for III–V substrate materials are not nearly so well developed as that for silicon. In general, they tend to produce non-uniform

layers that are not suitable for use as optical wave guides. However some researchers have been successful in making native-oxide waveguides on III-V materials. For example, Luo et al. [82] fabricated wet thermal native oxide layers on GaAlAs and used prism coupling measurements to study the dependence of the refractive index on aluminum concentration. They found that a sufficient Δn was produced to support waveguiding in the oxide layer. Deppe et al. have reported making low threshold vertical cavity lasers based on native-oxidation of AlAs [83].

4.7 Methods for Fabricating Channel Waveguides

The starting point for the fabrication of many types of channel waveguides is a substrate on which a planar waveguide has been formed by any of the methods previously described. The lateral dimensions of a number of different channel waveguides can then simultaneously be determined on the surface of the wafer by using standard photolithographic techniques, such as are used in the fabrication of electrical integrated circuits.

4.7.1 Ridged Waveguides Formed by Etching

The usual method of fabricating a ridged waveguide is to coat the planar waveguide sample with photoresist, expose the resist to uv light or x-rays through a contact printing mask that defines the waveguide shape, then develop the resist to form a pattern on the surface of the sample, as shown in Fig. 4.15 [84]. The photoresist [85, 86] is an adequate mask for either wet-chemical [87] or ion-beam-sputter etching [88]. Ion-beam-sputter etching, or *micro-machining* as it is sometimes called, produces smoother edges, particularly on curves, but it also causes some lattice

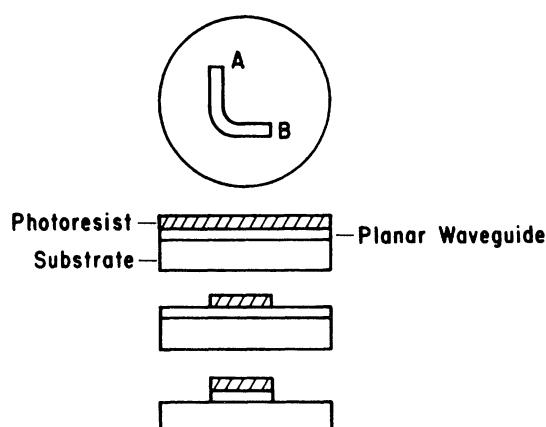
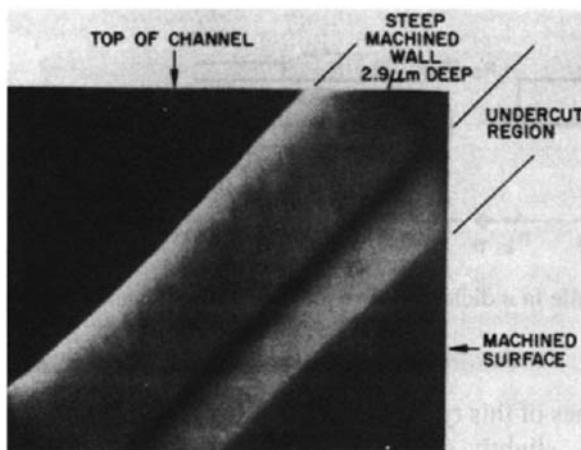


Fig. 4.15 Photoresist masking technique for channel waveguide fabrication

Fig.4.16 Ion beam etched channel waveguide. (Photo courtesy of Hughes Research Laboratories, Malibu, CA)

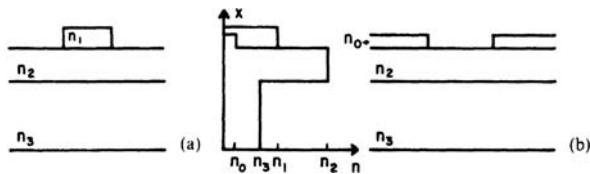


damage which must be removed by annealing if minimum optical losses are desired. A photograph of a ridged channel waveguide formed in GaAs by ion-beam micromachining is shown in Fig. 4.16 [88]. Photoresist makes a surprisingly good mask against bombarding ions because the mass of its organic molecules is so much less than the mass of Ar^+ or Kr^+ that there is relatively little energy transfer, as compared to that transferred in an Ar-GaAs interaction. For example, the difference in sputtering yields under 2 kV Ar^+ ion bombardment is such that a 1 μm thick photoresist mask allows the removal of 5 μm of GaAs in unmasked regions. Residual lattice damage after ion machining of GaAs can be removed by annealing the sample for about a half hour at 250°C.

Wet chemical etching can also be used in conjunction with a photoresist mask to define channel waveguide structures. Chemical etching produces no lattice damage, but it is very difficult to control the etch depth and profile. Most etchants are preferential with regard to crystal orientation, thus leading to ragged edges on curved sections of waveguides. However, preferential chemical etch effects can sometimes be used advantageously, as is done to define the rectangular mesas of semiconductor lasers [89].

An interesting method that appears to combine the advantages of ion-beam machining with those of wet chemical etching is an ion-bombardment-enhanced etching technique applied to LiNbO_3 by Kawabe et al.[90]. They produced 2 μm wide ridged waveguides in a Ti-diffused planar waveguiding layer by using a poly (methylmethacrylate) resist mask and electron beam lithography. The etching was accomplished by first bombarding the sample with a dose of $3 \times 10^{15} \text{ cm}^{-2}$ 60 keV Ar^+ at room temperature. Because of the relatively large energy of the Ar^+ ions, much lattice damage was generated but little sputtering of surface atoms occurred. The bombardment was followed by wet etching in diluted HF, which preferentially removes the damaged layer [91]. For 60 keV Ar^+ , the depth removed is 700 Å, but the bombardment per etch can be repeated a number of times to produce deeper

Fig. 4.17 a, b. Index of refraction profile in **a** dielectric strip loaded waveguide and **b** metal strip loaded waveguide



etching. Ion-bombardment-enhanced etching combines more accurate pattern definition of ion-beam micro-machining with the damage-free crystal surface that is obtained with wet chemical etching.

4.7.2 Strip-Loaded Waveguides

Instead of etching the lateral walls of a channel waveguide into a previously formed planar waveguide structure, a loading strip of dielectric can be deposited on the top surface to provide lateral confinement, as described in Section 3.2.2. The shape of this strip can be defined by the same photolithographic process that is used to make ridged waveguides, as described in Section 4.7.1. Again, either ion beam or wet chemical etching can be used in conjunction with the photoresist masking. Blum et al. [91] have made optical striplines of this type in GaAs by using a heavily doped GaAs loading strip, with index n_1 , slightly smaller than that in the waveguiding layer n_2 , as shown in Fig. 4.17. They used a wet chemical etch of NaOH and H_2O_2 , with photoresist masking. Also shown in Fig. 4.17 is a metal strip-loaded waveguide, with index n_0 in the metal. This type of waveguide can also be fabricated by photoresist masking and either ion-beam or chemical etching.

4.7.3 Masked Ion Implantation, Diffusion or Ion Exchange

Not all techniques for fabricating channel waveguides begin with a planar waveguide structure. Instead, channel wave guides can be formed by either diffusion [92], ion implantation [93] or ion exchange [94] of suitable dopant atoms directly into the substrate, but through a mask. Such waveguides are often called *buried* channel waveguides since they lie beneath the surface. Photoresist is not an effective mask in this case, because it cannot withstand the high temperature required for diffusion, nor does it have enough mass to block incident high-energy ions. Usually, deposited oxides, such as SiO_2 or Al_2O_3 , are used as diffusion masks, while noble metals such as Au or Pt are used as implantation masks. Photoresist is used, however, to define the pattern on the masking oxide or metal layer.

When waveguides are directly implanted or diffused into the substrate, the optical quality of the substrate material is critically important. It must have low optical absorption losses, as well as a smooth surface to avoid scattering. One of the most significant advantages of buried channel waveguide production by implantation or

diffusion through a mask is that it is a planar process; the surface of the OIC is not disrupted by ridges or valleys, so that optical coupling into and out of the OIC is facilitated, and problems of surface contamination by dust or moisture are minimized. Another advantage of this process is that it is less elaborate than those described in previous sections, because there is no need to fabricate a planar waveguide structure over the whole OIC chip before defining the strip waveguides. In order to define masked patterns with dimensions on the order of $1\text{ }\mu\text{m}$, electron beam lithography can be used [95, 96]. Also higher definition etched patterns can be obtained by using reactive ion etching [97, 98].

Both ion-implantation and ion-etching processes are usually done in conjunction with masks to define the lateral geometry of the waveguide. However state-of-the-art focused ion-beam systems are capable of forming beam diameters on the order of 50 nm, with a current density of 1 A/cm^2 . Using these systems, unique processing techniques such as maskless ion implantation and etching have been developed [99].

4.7.4 Focused Beam Writing Techniques

Rather than defining the geometry of a channel waveguide by conventional lithographic masking and etching (either dry or wet), it is also possible to define the guide by directly writing with a focused beam. This direct writing can take a number of different forms. Sure et al. [100] have used electron beam writing in conjunction with grayscale mask lithography of high-energy beam-sensitive glass, which is then used to create 3D structures into photoresist. Inductively coupled plasma (ICP) etching is then used to transfer the image to a silicon substrate. Focused ion beams have also been used to pattern channel waveguides. Arrand et al. [101] used a focused beam of nitrogen ions to fabricate waveguides in porous silicon (PSi). Teo et al. [102] report a direct-write technique for fabrication of 3D silicon waveguides that is compatible with standard Full Isolation by Porous Oxidized Silicon (FIPOS) technology [103]. In this method a focused beam of 250 keV protons is used to selectively slow down the rate of porous silicon formation during subsequent anodization, producing a silicon core surrounded by porous silicon cladding. The spot size of the beam in the reported experiments was on the order of 200 nm. The penetration depth (range) of 250 keV protons in silicon is approximately $2.5\text{ }\mu\text{m}$. Proton bombardment reduces the free-carrier density and increases the local resistivity of the material, as was explained in Section 4.3. During the subsequent electrochemical etching process, these defects act to trap holes from migrating to the silicon/electrolyte surface, reducing the rate of PSi formation in the irradiated regions. The end result is a silicon core surrounded by a region of porous silicon which has a lower index of refraction, as shown in Fig. 4.18.

Channel waveguides have also been fabricated by the use of focused laser beams to change the index of refraction. A comprehensive review of this technology, sometimes called laser “micromachining” is given by Gattass et al. [104]. Tightly focused femtosecond (fs) laser pulses are used to modify the refractive index of the substrate

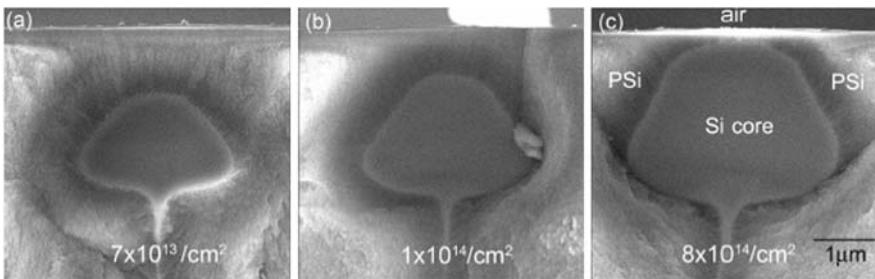


Fig. 4.18 Cross sectional SEM of the waveguides irradiated with a dose of (a) $7 \times 10^{13} / \text{cm}^2$, (b) $1 \times 10^{14} / \text{cm}^2$, and (c) $8 \times 10^{14} / \text{cm}^2$. [102]

material within the focal region of the laser beam, producing a waveguide inside the substrate. For example, Krol et al. [105] have used this method to fabricate channel waveguides inside two types of glasses: fused silica and Schon IOG-1, a phosphate glass. Fused silica (Coming 7940) is a common optical material ideal for passive device fabrication, while phosphate glasses, such as Schott IOG-I, are potentially well suited for fabricating active devices because of the high doping concentrations of rare earth ions allowed in the material. Waveguides were fabricated in these materials by focusing fs laser pulses from an amplified Ti-sapphire laser system (800 nm wavelength, 130 fs pulse duration, 1 kHz repetition rate) into the glass through a 10x microscope objective. Pulse energies typically ranged from 0.5 to 5 pJ. Single-mode waveguides fabricated in fused silica by this method exhibited losses less than 1 dB/cm at 633 nm wavelength.

The laser beam must be tightly focused in order to create a field strength great enough to produce the required laser-induced optical breakdown that ultimately results in the desired increase in index of refraction within the focal volume. For 800 nm laser pulses, the dominant mechanism is multiphonon absorption if the pulse length is less than 10 fs [106], while it is avalanche breakdown for pulse lengths of 30–200 fs [107]. Laser micromachining can be done with femtosecond pulses of just nanojoule energy as long as the repetition rate is kept high enough to accumulate the energy within the focal volume (typically 25 MHz or higher) [104]. Thus there is no need for an external device to amplify the laser pulse.

Problems

- 4.1 We wish to fabricate a planar waveguide for light of wavelength $\lambda_0 = 1.15 \mu\text{m}$ that will operate in the single (fundamental) mode. If we use the proton-bombardment, carrier-concentration-reduction method to form a $3 \mu\text{m}$ thick waveguide in GaAs, what are the minimum and maximum allowable carrier concentrations in the substrate? (Calculate for the two cases of p-type or n-type substrate material if that will result in different answers.)
- 4.2 If light of $\lambda_0 = 1.06 \mu\text{m}$ is used, what are the answers to Problem 4.1?

- 4.3 Compare the results of Problems 4.1 and 4.2 with those of Problems 2.1 and 2.2, note the unique wavelength dependence of the characteristics of the carrier-concentration-reduction type waveguide.
- 4.4 We wish to fabricate a planar waveguide for light of $\lambda_0 = 0.9 \mu\text{m}$ in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$. It will be a double layer structure on a GaAs substrate. The top (waveguiding) layer will be $3.0 \mu\text{m}$ thick with the composition $\text{Ga}_{0.9}\text{Al}_{0.1}\text{As}$. The lower (confining) layer will be $10 \mu\text{m}$ thick and have the composition $\text{Ga}_{0.17}\text{Al}_{0.83}\text{As}$. How many modes will this structure be capable of waveguiding?
- 4.5 What is the answer to Problem 4.4 if the wavelength is $\lambda_0 = 1.15 \mu\text{m}$?
- 4.6 A planar asymmetric waveguide is fabricated by depositing a $2.0 \mu\text{m}$ thick layer of Ta_2O_5 ($n = 2.09$) onto a quartz substrate ($n = 1.5$).
- How many modes can this waveguide support for light of 6328 \AA (vacuum wavelength)?
 - Approximately what angle does the ray representing the highest-order mode make with the surface of the waveguide?
- 4.7 A three-layer planar waveguide structure consists of a $5 \mu\text{m}$ thick waveguiding layer of GaAs with electron concentration $N = 1 \times 10^{14} \text{ cm}^{-3}$, covered on top with a confining layer of $\text{Ga}_{0.75}\text{Al}_{0.25}\text{As}$ with $N = 1 \times 10^{18} \text{ cm}^{-3}$, and on bottom with a confining layer of GaAs with $N = 1 \times 10^{18} \text{ cm}^{-3}$. If this waveguide is used to guide light of $\lambda_0 = 1.06 \mu\text{m}$, how many modes can it support?
- 4.8 For an asymmetric planar waveguide like that of Fig. 3.1 with $n_1 = 1.0$, $n_2 = 1.65$, $n_3 = 1.52$, $t_g = 1.18 \mu\text{m}$, and $\lambda_0 = 0.63 \mu\text{m}$, find the number of allowed TE modes.
- 4.9 If the waveguide of Prob. 4.4 were to be modified so as to support one additional mode, would it be practical to produce the required change in n by using the electro-optic effect? (Assume a uniform electric field to be applied over the thickness of the waveguide in the direction normal to the surface, with $r_{41} = 1.4 \times 10^{-12} \text{ m/V}$.)
- 4.10 Explain why the cutoff condition for a carrier-concentration-reduction type waveguide is independent of wavelength.
- 4.11 We desire an epitaxially-grown double-heterojunction $\text{In}_{(1-x)}\text{Ga}_x\text{As}_{(1-y)}\text{P}_y$ waveguide that will have a bandgap $E_g = 1.1\text{eV}$ in the waveguiding layer at room temperature.
- If lattice-matched layers are desired, what should be the relative concentrations of the constituent elements? i.e. What are x and y ?
 - What is the shortest wavelength that can be guided in this waveguide without excessive interband absorption?
- 4.12 We wish to make an epitaxially grown waveguide in $\text{In}_{(1-x)}\text{Ga}_x\text{As}_{(1-y)}\text{P}_y$ that has a bandgap of 0.9 eV at liquid nitrogen temperature (77 K). Find the values of x and y that are required.

References

1. P.K. Tien: *Appl. Opt.* **10**, 2395 (1971)
2. K.E. Wilson, E. Garmire, R.M. Silva, W.K. Stowell: *J. Opt. Soc. Am.* **71**, 1560 (1981)
3. R. Behrisch (ed.): *Sputtering by Particle Bombardment I*, Topics Appl. Phys., Vol. 47 (Springer, Berlin, Heidelberg 1981)
4. F. Zernike: Fabrication and measurement of passive components, in *Integrated Optics*, T. Tamir (ed.) 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979)
5. J.D. Plummer, J.D. Deal, P.B. Griffin: *Silicon VLSI Technology* (Prentice Hall, Upper Saddle River, NJ 2000) pp. 539–555
6. A.B. Glaser, G.E. Subak-Sharpe: *Integrated Circuit Engineering* (Addison-Wesley, Reading, MA 1977) pp. 169–181
7. B.J.H. Stadler, A. Gopinath: Magneto-optical garnet films made by reactive sputtering. *IEEE Transactions on Magnetics* **36**, 3957 (2000)
8. Y.B. Choi, S.J. Park, K.S. Shin, K.T. Jeong, S.H. Cho, D.C. Moon: The planar light waveguide type optical amplifier fabricated by sputtering method. APCC/OECC'99. Fifth Asia-Pacific Conference on Communications and Fourth Optoelectronics and Communications Conference, Volume: 2 (1999) pp. 1634–1635
9. T. Kitagawa, K. Hattori, Y. Hibino, Y. Ohmori: Neodymium-doped silica-based planar waveguide lasers. *IEEE J. Lightwave Technol.* **12**, 436 (1994)
10. M.P. Roe, M. Hempstead, J.L. Archambault, P.St.J. Russell, L. Dong: Strong photoinduced refractive index changes in RF-sputtered tantalum oxide planar waveguides. Conference on Lasers and Electro-Optics Europe (1994) p. 67
11. T.H. Hoekstra, P.V. Lambeck, H. Albers, T.J.A. Popma: Sputter-deposited erbium-doped yttrium oxide active optical waveguides. *Electron. Lett.* **29**, 581 (1993)
12. D. Ostrowsky, A. Jaques: *Appl. Phys. Lett.* **18**, 556 (1971)
13. K. Enbusu, M. Hikita, S. Tomaru, M. Usui, S. Imamura, T. Maruno: Multimode optical waveguide fabricated by UV curved epoxy resin for optical interconnection. APCC/OECC'99. Fifth Asia-Pacific Conference on Communications and Fourth Optoelectronics and Communications Conference, Vol.: 2 (1999) pp. 1648–1651
14. T. Sosnowski, H. Weber: *Appl. Phys. Lett.* **21**, 310 (1972)
15. D.A. Ramey, J.T. Boyd: *IEEE Trans. CAS* **26**, 1041 (1979)
16. R. Reuter, H. Franke, C. Feger: *Appl. Opt.* **27**, 4565 (1988)
17. T. Arakawa, T. Hasegawa, M. Kawaga: *Upt. Eng.* **42**, 898 (2003)
18. A. Neyer, T. Knoche, L. Müller: *Electron. Lett.* **29**, 399 (1993)
19. D.P. Prakash, D.V. Plant, D. Zhang, H.R. Fetterman: *SPIE Proc.* **1774**, 118 (1993)
20. P.K. Tien, G. Smolinsky, R. Martin: *Appl. Opt.* **11**, 637 (1972)
21. S.A. Campbell: *The Science and Engineering of Microelectronic Fabrication*, 2nd ed. (Oxford, New York 2001) pp. 39–65
22. J.L. Jackel, V. Ramaswamy, S.P. Lyman: *Appl. Phys. Lett.* **38**, 509 (1981)
23. Y. Liao, D. Chen, R. Lu, W. Wang: *Photo. Technol. Lett.* **8**, 548 (1996)
24. R. Twu, C. Huang, W. Wang: *Micro. Opti. Technol.* **48**, 2312 (2006)
25. J. Hukriede, D. Kip, E. Krataig: Photorefraction and thermal fixing in channel waveguides fabricated in lithium niobate by titanium and copper indiffusion. Conference Digest IEEE Conference on Lasers and Electro-Optics Europe (2000)
26. H.F. Taylor, W.E. Martin, D.B. Hall, V.N. Smiley: *Appl. Phys. Lett.* **21**, 325 (1972)
27. W.E. Martin, D.B. Hall: *Appl. Phys. Lett.* **21**, 325 (1972)
28. E.M. Zolatov, V.A. Kiselyov, A.M. Prokhorov, E.A. Sacherbakov: Determination of characteristics of diffused optical waveguides. OSA Topical Meeting on Integrated Optics, Salt Lake City, UT (1978)
29. B.L. Booth: Optical interconnection polymers, in *Polymers for Lightwave and Integrated Optics: Techniques and Applications*, L.A. Hornak (ed.) (Dekker, New York 1992) p. 232
30. T. Izawa, H. Nakagome: *Appl. Phys. Lett.* **21**, 584 (1972)

31. R.C. Alfernes: Titanium-diffused lithium niobate waveguide devices, in *Guided-Wave Optoelectronics*, T. Tamir (ed.), 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) pp. 145–206, in particular, p. 148
32. A. Tervonen, S. Honkanen, P. Poyhonen, M. Tahkokorpi: SPIE Proc. **1794**, 264 (1993)
33. P. Masalkar, V. Rao, R. Sirohi: SPIE Proc. **1794**, 271 (1993)
34. D.F. Geraghty, D. Provenzano, M. Morrell, S. Honkanen, A. Yariv, N. Peyghambarian: Ion-exchanged waveguide add/drop filter. Electron. Lett. **37**, 829 (2001)
35. B. Buchold, E. Voges: Planar arrayed-waveguide grating multi/demultiplexers based on ion-exchanged waveguides in glass. IEE Colloquium on WDM Technology and Applications (Digest No. 1997/036) (1997) pp. 10/1–10/5
36. D. Nir, S. Ruschin, A. Hardy, D. Brooks: Proton-exchanged periodically segmented channel waveguides in lithium niobate. Electron. Lett. **31**, 186 (1995)
37. H. Ryssel, H. Glawischning (eds.): *Ion Implantation*, Springer Ser. Electrophys., Vols. 10 and 11 (Springer, Berlin, Heidelberg 1982 and 1983)
38. W.S. Johnson, J.F. Gibbons: *Projected Range Statistics in Semiconductors* (Stanford Univ. Press, Standford, CA 1969)
39. R. Standley, W.M. Gibson, J.W. Rodgers: Appl. Opt. **11**, 1313 (1972)
40. E. Garmire, H. Stoll, A. Yariv, R.G. Hunsperger: Appl. Phys. Lett. **21**, 87 (1972)
41. M.A. Mentzer, R.G. Hunsperger, S. Sriram, J. Bartko, M.S. Wlodawski, J.M. Zavada, H.A. Jenkinson: Opt. Eng. **24**, 225 (1985)
42. M. Barnoski, R.G. Hunsperger, R. Wilson, G. Tangonan: J. Appl. Phys. **44**, 1925 (1973)
43. J. Zavada, H. Jenkinson, T. Gavanis, R.G. Hunsperger, M. Mentzer, D. Larson, J. Comas: SPIE Proc. **239**, 157 (1980)
44. J.P. Donnelley, A.G. Foyt, W.T. Lindley, G.W. Iseler: Solid State Electron, **13**, 755 (1970)
45. P. Bhattacharya: *Semiconductor Optoelectronic Devices* (Prentice-Hall, Englewood Cliffs, NJ 1944) pp. 133–137, 294–299
46. E. Garmire: Semiconductor components for monolithic applications, in *Integrated Optics*, T. Tamir (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979) Chap. 6, in particular, pp. 293–301
47. T. Moss, G. Hawkins: Infrared Phys. **1**, 111 (1961)
48. D. Hill: Phys. Rev. A **133**, 866 (1963)
49. J. Shah, B.I. Miller, A.E. DiGiovanni: J. Appl. Phys. **43**, 3436 (1972)
50. J.T. Boyd: IEEE J. **QE-8**, 788 (1972)
51. H.C. Casey Jr., D.D. Sell, M.B. Panish: Appl. Phys. Lett. **24**, 633 (1974)
52. V. Evtuhov, A. Yariv: IEEE Trans. MTT-**23**, 44 (1975)
53. Zn.I. Alferov, V.M. Andreev, E.L. Portnoi, M.K. Trukn: Sov. Phys. – Semiconductors **3**, 1107 (1970)
54. J.H. McFee, R.H. Nahory, M.A. Pollack, R.A. Logan: Appl. Phys. Lett. **23**, 571 (1973)
55. F.K. Reinhart, R.A. Logan, T.P. Lee: Appl. Phys. Lett. **24**, 270 (1974)
56. S.M. Jensen, M.K. Barnoski, R.G. Hunsperger, G.S. Kamath: J. Appl. Phys. **46**, 3547 (1975)
57. M.G. Crawford, W.O. Groves: IEEE Proc. **61**, 862 (1973)
58. S. Kamath: Epitaxial GaAs-(GaAl)As layers for integrated optics. OSA Topical Meeting on Integrated Optics, New Orleans, LA (1974)
59. H. Kressel (ed.): *Semiconductor Devices for Optical Communication*, 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982)
60. C.M. Wolfe, G.E. Stillman, M. Melngallis: Epitaxial growth of InGaAs-GaAs for integrated optics. OSA Topical Meeting on Integrated Optics, New Orleans, LA (1974)
61. M. Kawabe: Appl. Phys. Lett. **26**, 46 (1975)
62. K. Nakajimi, A. Yamaguchi, K. Akita, T. Kotani: J. Appl. Phys. **49**, 5944 (1979)
63. R.U. Martinelli: LEOS'88, Santa Clara, CA (1988) Digest p. 55
64. M.R.T. Pearson, P.E. Jessop, D.M. Bruce, S. Wallace, P. Mascher, J. Ojha: Fabrication of SiGe optical waveguides using VLSI processing techniques. IEEE J. Lightwave Tech. **19**, 363 (2001)

65. A.A. Chernov (ed.): *Modern Crystallography III, Crystal Growth*, Springer Ser. Solid-State Sci., Vol. 36 (Springer, Berlin, Heidelberg 1984)
66. K. Ploog, K. Graf: *Molecular Beam Epitaxy of III-V compounds, a Comprehensive Bibliography 1958–1983* (Springer, Berlin, Heidelberg 1984)
67. M.A. Herman, H. Sitter: *Molecular Beam Epitaxy*, 2nd edn., Springer Ser. Mater. Sci., Vol. 7 (Springer, Berlin, Heidelberg 1996)
68. W.T. Tsang: Appl. Phys. Lett. **38**, 587 (1981)
69. J.C.M. Hwang, J.V. DiLorenzo, P.E. Luscher, W.S. Knodel: Solid State Techn. **25**, 166 (1982)
70. C. Goldstein, C. Stark, J. Emory, F. Gaberit, D. Bonnevie, F. Poingt, M. Lambert: J. Crystal Growth **120**, 157 (1992)
71. R.G. Walker, R.C. Goodfellow: Electron. Lett. **19**, 590 (1983)
72. T. Matsumoto, P. Bhattacharya, M.J. Ludowise: Appl. Phys. Lett. **42**, 52 (1983)
73. H. Ishiguro, T. Kawabata, S. Koike: Appl. Phys. Lett. **51**, 12 (1987)
74. H. Jvergensen: Microelectron. Eng. **18**, 119 (1992)
75. A. Yariv: *Optical Electronics*, 4th edn. (Holt, Rinehart and Winston, New York 1991) pp. 309–316
76. B.G. Streetman: *Solid State Electronic Devices*, 3rd edn. (Prentice-Hall, Englewood Cliffs, NJ 1990) p. 147
77. S.A. Campbell: *The Science and Engineering of Microelectronic Fabrication*, 2nd ed. (Oxford, New York, 2001) pp. 68–95
78. C. Boulas, S. Valertte, E. Parrens, A. Fournier: Low loss multimode waveguides on silicon substrate. Electron. Lett. **28**, 1648 (1992)
79. Q. Lai, P. Pliska, J. Schmid, W. Hunziker, H. Melchior: Formation of optical slab waveguides using thermal oxidation of SiO₂. Electron. Lett. **29**, 1648 (1992)
80. A.V. Tomov, V.V. Filippov, V.P. Bondarenko: Pis'ma Zh. Tekh. Fiz. **23**, 86 (1997)
81. O.K. Sparacin, S.J. Spector, L.C. Kimerling: J. Lightwave Technol. **23**, 2455 (2005)
82. Y. Luo, D.C. Hall, L. Kou, O. Bium, H. Hou, L. Steingart, J.H. Jackson: Optical properties of AlGaAs heterostructure native oxide planar waveguides LEOS'99. IEEE
83. D.G. Deppe, D.L. Huffaker, H. Deng, C.C. Lin: Engineering Foundation Conference on High Speed Optoelectronic Devices for Communications and Interconnects, San Luis Obispo, CA (1994)
84. E. Spiller, R. Feder: X-ray lithography, in *X-Ray Optics*, H.-J. Queiser (ed.), Topics Appl. Phys., Vol. 22 (Springer, Berlin, Heidelberg 1977)
85. R.A. Bartolini: Photoresists, in *Holographic Recording Materials*, ed. by H.M. Smith, Topics Appl. Phys., Vol. 20 (Springer, Berlin, Heidelberg 1977)
86. H.I. Bjelkhagen: *Silver-Halide Recording Materials for Holography and Their Processing*, 2nd edn., Springer Ser. Opt. Sci., Vol. 66 (Springer, Berlin, Heidelberg 1995)
87. M.C. Rowland: The preparation and properties of gallium arsenide, in *Gallium Arsenide Lasers*, C.H. Gooch (ed.) (Wiley-Interscience, New York 1969) p. 166
88. S. Somekh, E. Garmire, A. Yariv, H. Garvin, R.G. Hunsperger: Appl. Opt. **12**, 455 (1973)
89. A.R. Goodwin, D.H. Lovelace, P.R. Selway: Opto-Electron. **4**, 311 (1972)
90. M. Kawabe, S. Hirata, S. Namba: IEEE Trans. CAS-**26**, 1109 (1979)
91. F.A. Blum, D.W. Shaw, W.C. Holton: Appl. Phys. Lett. **25**, 116 (1974)
92. H.F. Taylor, W.E. Martin, D.B. Hall, V.N. Smiley: Appl. Phys. Lett. **21**, 95 (1972)
93. S. Somekh, E. Garmire, A. Yariv, H. Garvin, R.G. Hunsperger: Appl. Opt. **13**, 327 (1974)
94. G. Li, K.A. Winick, H.C. Griffin, J.S. Hayden: Appl. Opt. **45**, 1743 (2006)
95. S.M. Sze: *VLSI Technology*, 2nd edn. (McGraw-Hill, New York 1988)
96. D.F. Barbe: *Very Large Scale Integration (VLSI)*. 2nd edn., Springer Ser. Electrophys., Vol. 5 (Springer, Berlin, Heidelberg 1982)
97. J.L. Jackel, R.E. Howard, E.L. Hu, S.P. Lyman: Appl. Phys. Lett. **38**, 907 (1981)
98. M.A. Bosch, L.A. Coldren, E. Good: Appl. Phys. Lett. **38**, 264 (1981)
99. K. Gamo: Mater. Sci. Eng. B: Solid-State Mater. for Adv. Techn. B **9**, 307 (1991)
100. A. Sure, T. Dillon, J. Murakowski, C. Lin, D. Pustai and D.W. Prather: Opt. Express **11**, 3555 (2003)

101. H.F. Arrand, T.M. Benson, P. Sewell, A. Loni: *J. Lumin.* **80**, 199 (1999)
102. E.J. Teo, A.A. Bettoli, M.B.H. Breese1, P. Yang, G.Z. Mashanovich, W.R. Headley, G.T. Reed, D.J. Blackwood: *Optics Express* **16**, 573 (2008)
103. K. Imai: *Solid State Electron.* **24** 150 (1981)
104. R.R. Gattass, L.R. Cerami, E. Mazur: *Proc. Int. Workshop on Optical and Electronic Device Technology for Access Network*, San Jose, CA, 51 (2005)
105. D.M. Krol, J.W. Chen, T. Huser, S.H. Risbud, J. Hayden: *CLEO/Europe 2003 Conference on Lasers and Electro-Optics Europe*, **346** (2003)
106. M. Lenzner, J. Kruger, S. Sartania, Z. Cheng, C. Spielmann, G. Mourou, W. Kautek, F. Ksausz: *Physical Review Letters* **80**, 4076 (1998)
107. C.B. Schaffer, A. Brodeur, E. Mazur: *Meas. Sci. Technol.* **12** 1784 (2001)

Chapter 5

Polymer and Fiber Integrated Optics

In recent years, the interest in making waveguides and other integrated optic devices in polymer materials has grown rapidly. The driving force behind this development is cost reduction. Semiconductor materials and dielectric materials such as lithium niobate are relatively expensive and the processes used to fabricate devices in those materials are very complex. As the size and complexity of fiber optic telecom and datacom systems has grown, leading to fiber connections to the office and home, the demand for large quantities of inexpensive integrated optic devices has increased. As a result, research has been directed toward producing in polymers many of the devices that formerly were made only in III-V semiconductors or in lithium niobate. It also has been demonstrated that some integrated optic devices can be made from glass or plastic optical fiber waveguides as well as polymers. An example of this latter type of device is the array waveguide (AWG), which performs a spatial dispersion of different wavelengths similar to that produced by a prism [1].

5.1 Types of Polymers

There are a number of different polymers that can be considered for use in optical integrated circuits. Each one of these has a different set of optical, electrical and mechanical properties. Some polymers create glassy films, while others produce flexible films. The key optical properties are the index of refraction and the attenuation or optical loss, which is usually measured in dB/cm. These are both functions of wavelength and thus must be determined at the wavelengths of interest. Other properties of interest are thermal and mechanical stability and any possible birefringence. The glass transition temperature is also an important parameter. It is the temperature at which the polymer relaxes to a stable form. The layers formed by polymers can be either crystalline or amorphous. Polymers are available commercially in bulk quantities from a number of different suppliers. The key properties of some of the more commonly used optical polymers have been summarized by Eldada and Shaklette [2] as shown in Table 5.1. It can be seen that optical losses are generally on the order of tenths of a dB per centimeter at the wavelengths corresponding to the first, second and third “telecommunications windows”, which are

Table 5.1 Key properties of optical polymers developed worldwide [2]© 2000 IEEE

Manufacturer	Polymer Type [Trade Name]	Patterning Techniques	Optical Loss, dB/cm [at wavelength, nm]	Other Properties [at wavelength, nm]
Allied Signal	Acrylate	Photoexposure/wet etch, RIE, laser ablation	0.02 [840] 0.2 [1300] 0.5 [1550]	Birefringence: 0.0002 [1550] Crosslinked, T_g : 25°C Environmentally stable
	Halogenated Acrylate	Photoexposure/wet etch, RIE, laser ablation	< 0.01 [840] 0.03 [1300] 0.07 [1550]	Birefringence: < 0.000001 [1550] Crosslinked, T_g : -50°C Environmentally stable
	Fluorinated Polyimide [Ultradel™]	Photoexposure/wet etch	0.4 [1300] 1.0 [1550]	Birefringence: 0.025 Crosslinked, Thermally stable
Amoco	Benzocyclobutene [Cyclotene™]	RIE	0.8 [1300] 1.5 [1550]	T_g : > 350°C
	Perfluorocyclobutene [XU 35121]	Photoexposure/wet etch	0.25 [1300] 0.25 [1550]	T_g : 400°C
Dow Chemical	Acrylate [Polyguide™]	Photolocking	0.18 [800] 0.2 [1300] 0.6 [1550]	Laminated sheets Excimer laser machinable
DuPont	Polyetherimide [Ultem™]	RIE, laser ablation	0.24 [830]	Thermally stable
	PMMA copolymer [P2ANS]	Photobleaching RIE	1.0 [1300] 0.6 [1550]	NLO polymer Thermally stable
	Halogenated Acrylate	RIE	0.02 [830] 0.07 [1310] 1.7 [1550]	Birefringence: 0.000006 [1310] T_g : 110°C
General Electric	Deuterated Polysiloxane	RIE	0.17 [1310] 0.43 [1550]	Environmentally stable
	Fluorinated Polyimide	RIE	TE: 0.3, TM: 0.7 [1310]	Environmentally stable

840, 1300 and 1500 nm, respectively. A number of different technologies have been used to pattern the polymer layers after they have been deposited on the substrate, including photolithography or e-beam lithography followed by wet etching, reactive ion etching (RIE) and laser ablation.

Polymers are currently available as either homopolymers or block copolymers. A homopolymer is a single chemical structure consisting of identical chemical units that are covalently bonded to one another to form a linear polymer chain. A block copolymer contains two polymers covalently bonded together. When heated above their glass transition temperature, block copolymers can separate into stratified layers of the two constituent polymers, suggesting that a self-forming, multi-layer structure may be a possibility.

5.2 Polymer Processing

The processing of polymers to create thin films such as those used in waveguides or integrated optic devices is much less complex, and therefore much less expensive, than the processes that have been described in Chapter 4 for III–V semiconductors and for lithium niobate. Generally the polymer is dissolved in a solvent and then spun onto the substrate wafer using a conventional photoresist spinner. The polymer layer is then heated above its glass transition temperature for a period of time to drive off any remaining solvent and to relax any strain. The optical and mechanical properties of the polymer film are dependent upon the process variables such as polymer concentration in the solvent, molecular weight of the polymer, spin speed, anneal time and temperature. For example, the molecular weight will affect the viscosity of the polymer in solution, which in turn will affect the thickness of the resultant polymer layer.

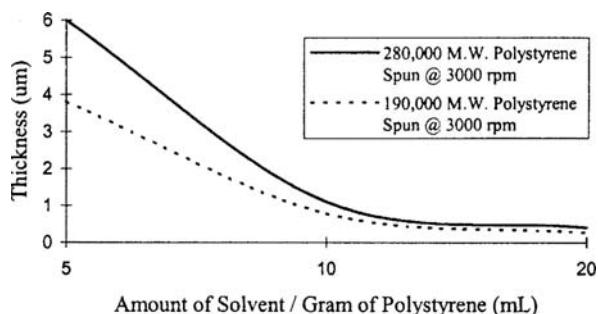
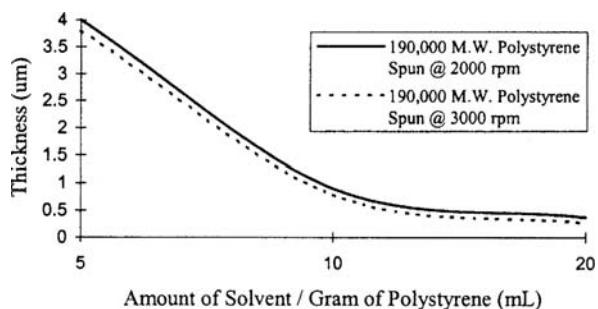
5.2.1 Processing of Polystyrene

Polystyrene homopolymer characteristics as functions of processing variables have been studied by Everhart [3]. Polystyrene is an attractive material for waveguide fabrication because it has an index of refraction of 1.5894 (somewhat larger than glass) a glass transition temperature of 100°C and is transparent in the visible and near-infrared wavelength range. Samples of two different molecular weights were studied, 280 000 and 190 000. These were dissolved in toluene to form solutions of several different concentrations, which were then spun onto glass substrates to form polymer films. The thickness of the resulting film depended on molecular weight, concentration, spin speed, spin time, anneal temperature and anneal time as shown in Table 5.2.

The thickness of the polystyrene film is seen to range from approximately 2600 to 59 900 angstroms, a range that is appropriate for waveguides and other integrated optic devices. The thickness of the polymer layers depends most strongly on the

Table 5.2 Dependence of film thickness on process variables [3]

Sample	Molecular Weight	Conc. (ml/gram)	Spin Speed (rpm)	Spin Time (sec.)	Anneal Temp. (°C)	Anneal Time (min.)	Average Thickness (angstroms)
1	280 000	5	3000	60	110	45	59932
2	280 000	10	3000	40	110	45	10710
3	280 000	20	3000	40	110	45	4084
4	190 000	5	2000	60	110	45	40297
5	190 000	10	2000	40	110	45	9276
6	190 000	20	2000	40	110	45	3797
7	190 000	5	3000	60	110	45	38016
8	190 000	10	3000	40	110	45	7848
9	190 000	20	3000	40	110	45	2598

Fig. 5.1 Effect of molecular weight on thickness [3]**Fig. 5.2** Effect of spin speed on thickness [3]

concentration of the bulk polymer that is dissolved in the solvent, as can be seen in Figs. 5.1 and 5.2, with a lesser dependence on the other process variables. The dependence on different variables is as one would intuitively expect. For example, larger molecular weight produces a thicker layer and faster spin speed yields a thinner layer.

Table 5.3 Process for polyimide channel waveguide fabrication [7]

-
1. Clean substrate with acetone, methanol and isopropyl alcohol.
 2. Spin coat polyimide, typically 10 s at 500 rpm.
 3. Soft bake at 120°C 20 or 30 min.
 4. Apply photoresist by spinning on.
 5. Bake at 90°C for 10 min.
 6. Expose photoresist using mask aligner.
 7. Develop photoresist (If the photoresist contains an alkali the polyimide will also be etched in this step.)
 8. Etch polyimide. (Use either an alkali wet etch or a sputter dry-etching process.)
-

Polystyrene films of this type with thickness in the range of 1–1.2 μm formed waveguides with relatively low losses of approximately 1 dB/cm.

5.2.2 Processing of Polyimide

Polyimides are polymers with excellent thermal characteristics, resisting extended exposure to temperatures as high as 350°C. They also have good solvent resistance and electrical properties. Since they are transparent to visible and near-infrared wavelengths they are good candidates as waveguide materials. In fact, a number of researchers have reported making waveguides in polyimide [4–7]. Polyimide can be dissolved in a solvent and spun onto the substrate in a fashion similar to that used for polystyrene. A typical process, described by Zhang [7], is shown in Table 5.3.

The above process can be used for either glass or silicon substrates. However, since silicon has an index of refraction that is greater than that of polyimide, it is necessary to first deposit or grow a layer of silicon dioxide ($n = 1.46$) as a light confining layer before spinning on the polyimide. In the case of glass substrates the polyimide can be directly applied because its index of refraction (typically 1.56–1.7) is larger than that of glass (1.5).

5.2.3 Post-Deposition Processing

Standard photolithographic techniques [8] are generally used to pattern the waveguide's lateral dimensions. However, direct writing with lasers or e-beams is also possible [9]. Polymers can be etched with solvents [10, 11] or by Reactive Ion Etching (RIE) [12, 13]. Losses in etched waveguides typically range from 0.1 dB/cm at 850 nm wavelength to 0.3 dB/cm at 1300 nm. Loss of 0.3 dB/cm at 1300 nm wavelength has also been observed in single-mode polymer waveguides fabricated by injection molding of polymethylmethacrylate (PMMA) substrates, with waveguide grooves (4–7 μm wide) filled with fully deuterated ethyleneglycoldimethacrylate (EGDMA) [14].

5.3 Applications of Polymer Waveguide Interconnections

Polymer waveguides have been used in some fairly sophisticated optical signal transmission applications. For example, Prakash et al. [15] employed polyimide channel waveguides fabricated using direct laser writing as the interconnection media between 50 GHz modulated sources and high-speed photodetectors. They suggest that such interconnections could potentially be used in the implementation of sophisticated wafer-scale systems such as phased array radars.

The potential of polymer waveguides as interconnections between devices and/or circuit boards is also demonstrated by the eight-channel transmit/receive link fabricated by Thomson et al. [16]. This demonstration link involved a fiberglass epoxy (FR4) photonic-circuit board containing eight-channel transmitter and receiver packages, interconnected by DuPont Polyguide TM [17] waveguides, as shown in Fig. 5.3. The waveguide circuits include eight-channel arrays of straights, cross-throughs, curves and self-aligning interconnects to multi-fiber ribbon at the edge of the board. The Polyguide waveguides were coupled to the TX and RX chips by

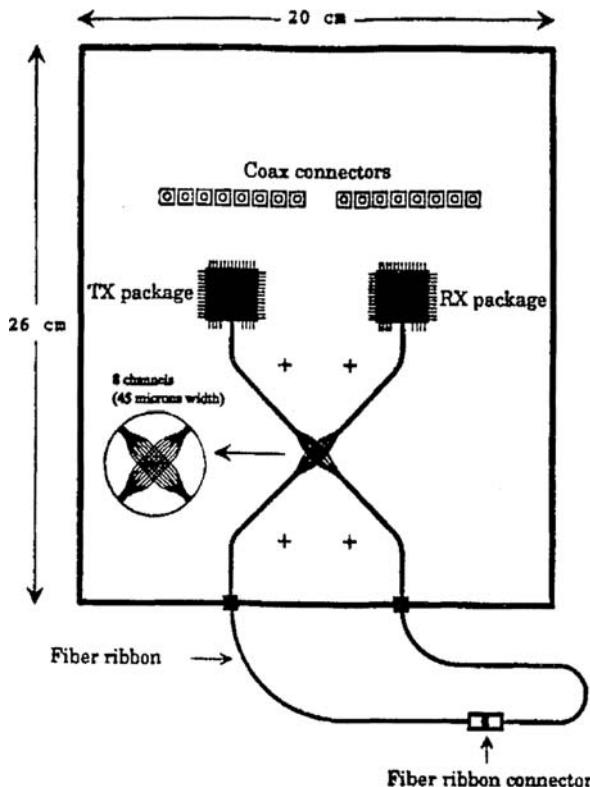


Fig. 5.3 Diagram of photonic circuit board 8-channel link demonstrator [16]

means of 45° out-of-plane reflecting mirrors formed on the ends of the waveguides beneath the chips (see Chapter 7 for more details of this coupling method.) To minimize optical losses in the curved and diagonal waveguides, the masks were made with $0.1/\mu\text{m}$ resolution. Data transmission at up to 1.25 Gbit/s per channel was achieved with this demonstration board with a bit error ratio less than 10^{-11} . Adjacent channel crosstalk was -31 dB, including electrical crosstalk on the receiver array chip and optical crosstalk.

A variety of photochemically set optically transparent polymers based on combinations of acrylate monomers/oligomers has been reported by Eldada et al. [18]. On photochemical exposure these monomer systems form highly crosslinked networks that have low absorption in the range from 400 to 1600 nm. When fully

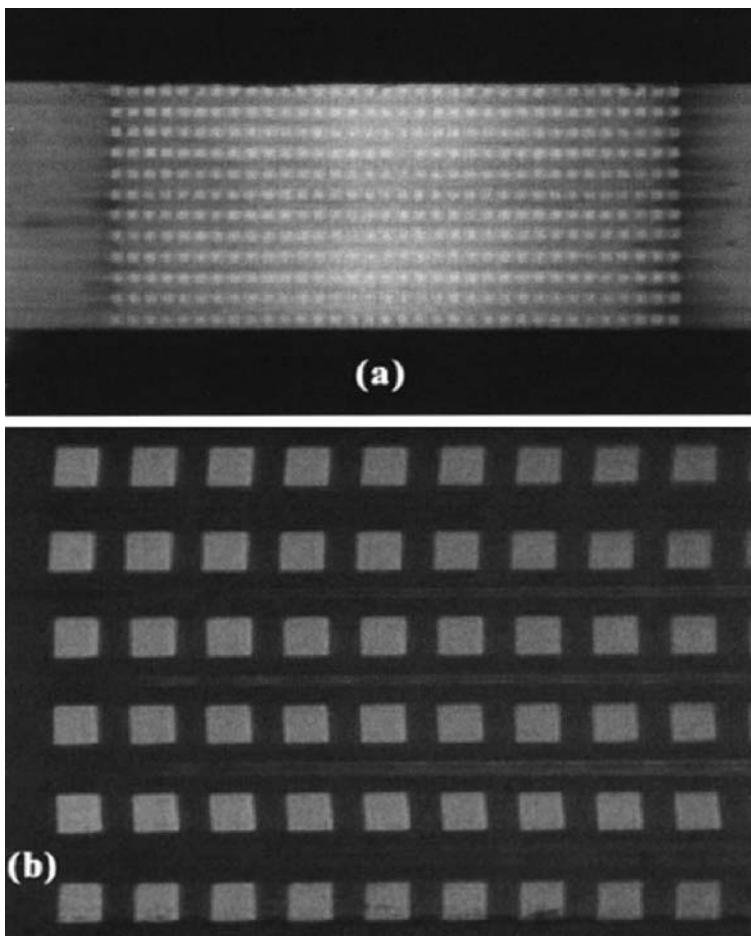


Fig. 5.4 a,b A twelve-layer polymer interconnect for a two-dimensional VCSEL array [19] **a** entire array 1.5 mm height (12 layers at $125 \mu\text{m}$ each), 4.25 mm width **b** close up view

polymerized, these polymers have an index of refraction in the range from 1.3 to 1.6, depending on the exact mixture of polymers. Waveguides are formed photolithographically, with the liquid monomer mixture polymerizing when it is illuminated by UV light by either exposure through a mask or by direct laser writing. The polymers can be deposited on a variety of substrates, including both rigid and flexible types.

One of the key advantages of polymer waveguides is that they can be deposited on flexible plastic substrates. Thus they can be used for backplane "wiring" or for planar strip cables to connect to arrays of light emitters or detectors. An example of this type of application is the high-density interconnect cable for two-dimensional VCSEL (vertical cavity surface emitting laser) arrays described by Gallo et al. [19]. Multimode polymer waveguide ribbons 34-channels wide were stacked in arrays 12-layers high (34×12 channels) with center-to-center waveguide spacing of 125 microns between layers and 90 microns within a layer. No measurable crosstalk between channels was observed even when separation between multimode waveguides was reduced to a 4-micron gap. These flexible polymers provide an out-of-plane bend radius of less than 5 mm that simplifies VCSEL packaging requirements and volume. A photograph of a stacked waveguide array is shown in Fig. 5.4.

Transitioning the waveguide pitch within and between the polymer layers from 125 to 250 microns enables interfacing of high-density VCSEL array to standard fiber ribbons. Passive fiber pigtailing to 62.5/125 fibers was achieved with < 0.5 dB loss. Pigtailing can be avoided entirely by direct connectorization of the polymer waveguide arrays with industry standard MT connectors as shown in Fig. 5.5.

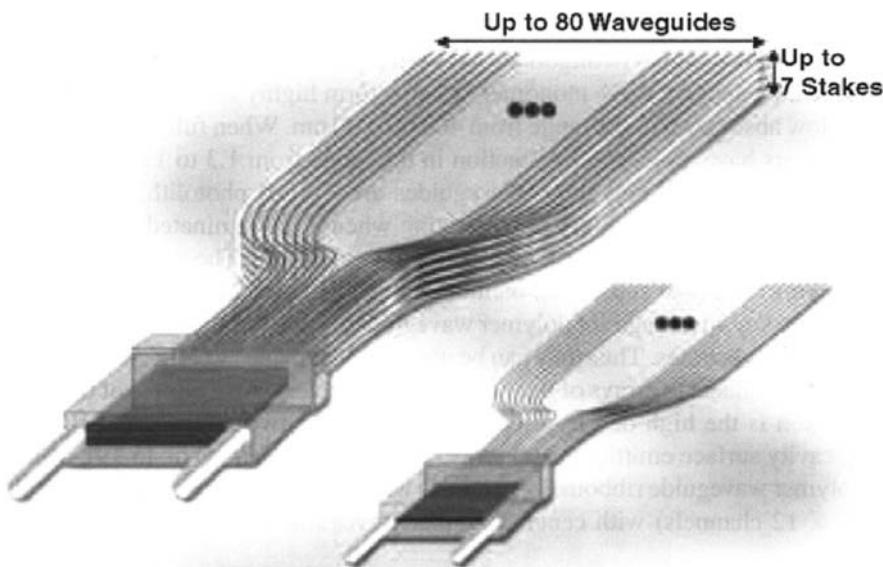


Fig. 5.5 Five layer polymer array packaged in an MT ferrule [19]

5.4 Polymer Waveguide Devices

In addition to being used as an interconnecting media, polymer waveguides can form the basic elements for a variety of both passive and active optoelectronic devices. In cases in which a polymer device can be substituted for one made in a III-V semiconductor or in lithium niobate without undue loss of performance, a great saving of cost can usually be obtained. This cost saving comes about not only because of reduced materials expense but also because the processing of polymers is much simpler than that of semiconductors or crystalline dielectrics such as lithium niobate. The more prominent devices that have been fabricated in polymers are listed in Table 5.4. They have been divided into two categories: *passive*, in which the properties of the devices are constant and operation depends on coupling, reflection, diffraction, refraction, absorption, etc.; and *active*, in which the properties are altered by application of a voltage, current, force or chemical reaction to produce the desired effect.

5.4.1 Passive Polymer Devices

Branching couplers and splitters can be easily fabricated in polymers by beginning with a planar polymer waveguiding layer and defining a branching pattern by means of the standard photolithographic techniques that were described in Chapter 4 to produce either etched ridged waveguide structures or buried channel waveguide structures. Van der Linden et al. have used buried channel waveguide 3 dB (50/59) splitters to make a compact multichannel in-line optical power meter [20], while etched polymer waveguides have been used by Pliska et al. [21] to produce nonlinear optical devices. It also is possible to fabricate buried branching waveguide structures by direct laser writing of the patterns into the polymer [22] or by UV exposure through a photomask [18]. This latter process is diagrammed in Fig. 5.6.

The polymers used are similar to a negative photoresist and can be applied to the substrate by spin-coating. Substrates such as glass, quartz, silicon, flexible plastic films or glass-filled epoxy printed circuit boards can be used. Exposure is accomplished with a Hg or Hg-Xe arc lamp through a photomask. Development is done with a conventional organic solvent such as methanol. If desired, an over-cladding

Table 5.4 Polymer waveguide devices

(Passive)	(Active)
Coupler/splitter	Light emitter
Grating/filter	LASER
Delay line	Light detector
Mode converter	Optical modulator/switch
Attenuator	Accelerometer
Array waveguide (AWG)	Chemical detector

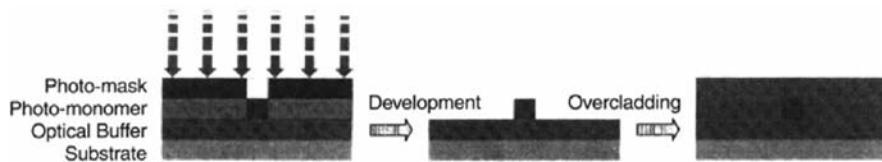


Fig. 5.6 Mask-based process for making photopolymer waveguides [18]

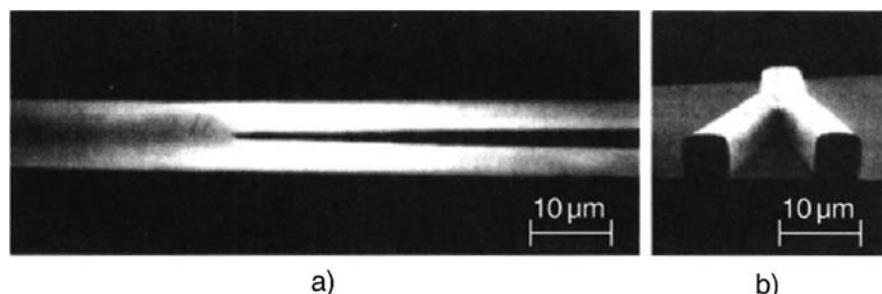


Fig. 5.7 a, b Photomicrograph of 1×2 splitter **a** top view **b** cross-sectional view [18]

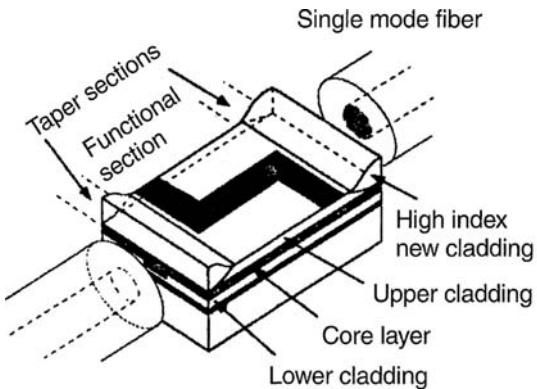
layer of a material such as SiO_2 can be added to protect the polymer device. A photograph of a 1×2 splitter is shown in Fig. 5.7.

In addition to making branching splitters and couplers it is also possible to fabricate mode converters and mode size transformers in polymers by varying the thickness of the waveguide. Chen et al. [23] have used that approach to produce a tapered mode expander that allows the mode of a rectangular waveguide to be more efficiently coupled to a cylindrical optical fiber. The taper converts the highly elliptical waveguide mode to a larger, more circular shape that more closely matches that of the fiber core. The taper, which is 0.5–2 mm in length, consists of a high index material that allows the effective thickness of the waveguide to increase at its ends. Reactive ion etching (RIE) techniques are used to make the taper employing either a tapered photoresist mask or a shadow masking technique.

The tapered regions are placed outside of the metallic contact electrode as shown in Fig. 5.8 to avoid the losses associated with it and insure that there is no trade-off between mode size and propagation loss. The performance of this mode size transformer was found to be independent of both waveguide width and optical beam polarization, and a 1.6 dB reduction in coupling loss was observed at each end.

As a final example of passive polymer devices, consider the case of the array waveguide (AWG). This is a device that performs the spatial dispersion of light of different wavelengths, producing an output that is similar to that of a prism. However, when fabricated in a planar polymer waveguide integrated circuit, it is easier to couple to optical fiber input and output channels and is less sensitive to vibration than is a prism. Because of its simplicity and effectiveness the AWG is widely used as a multiplexer/demultiplexer in dense wavelength division multiplexing (DWDM)

Fig. 5.8 Diagram of mode size transformer (Length not drawn to scale) [23]



applications. A thorough explanation of the principles of operation and fundamental characteristics of the AWG are given by Kaneko et al. [24]. A diagram of an AWG demultiplexer is shown in Fig. 5.9, in which a beam composed of N_{ch} channels of wavelength division multiplexed light enters on the left. However, the device is reciprocal and hence also could function as a multiplexer if N_{ch} channels of different wavelengths were introduced with light propagating from right to left.

The AWG consists of input and output waveguides, two focusing slab regions and a phased array of waveguides with a constant path length difference, ΔL , between neighboring waveguides. In the first slab region, the input waveguide separation is

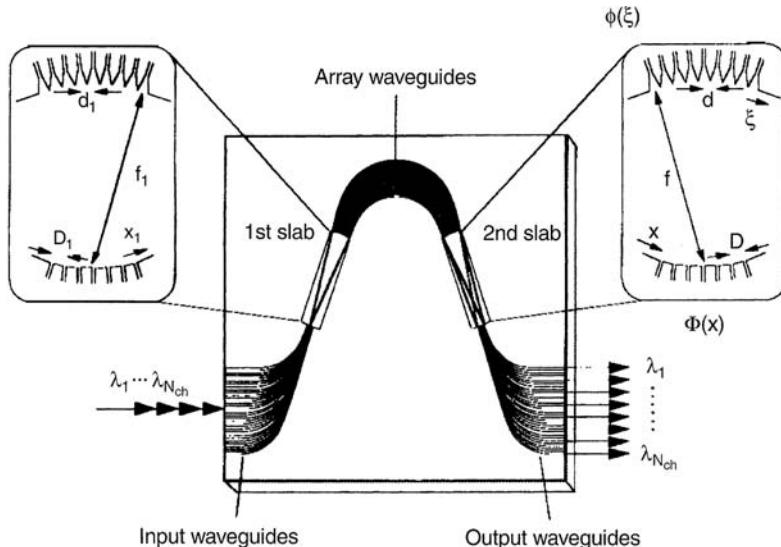


Fig. 5.9 Array waveguide demultiplexer (or multiplexer by reciprocity) [24] © 1999 IEEE

D_1 , the array waveguide separation is d_1 , and the radius of curvature is f_1 . In the second slab region, the output waveguide separation is D_2 , the array waveguide separation is d_2 , and the radius or curvature is f_2 . Usually the waveguide parameters are the same on the input and output sides, so that $D_1 = D$, etc. In that case, it can be shown that the required path length difference between adjacent waveguides is given by [24]:

$$\Delta L = (n_s d D \lambda_0) / (N_{\text{ch}} f \Delta \lambda), \quad (5.1)$$

where n_s is the index of refraction in the slab regions, λ_0 is the center wavelength of the WDM system, and $\Delta\lambda$ is the wavelength spacing between adjacent channels. For maintaining the uniqueness of each channel and avoiding crosstalk it is essential that only the first order of the AWG spatial dispersion be used, and that limits its spatial range and number of possible channels. The spatial separation of the m th and $(m + 1)$ th focused beams for the same wavelength is given by:

$$X_{\text{FSR}} = \lambda_0 f / n_s d. \quad (5.2)$$

The number of available wavelength channels is then:

$$N_{\text{ch}} = X_{\text{FSR}} / D = \lambda_0 f / n_s d D. \quad (5.3)$$

Array waveguides have been fabricated for hundreds of channels with no crosstalk problems. In Chapter 20 the application of AWGs to optical add/drop multiplexers (OADM) will be discussed. These have become very important elements of DWDM telecom and datacom systems.

5.4.2 Active Polymer Devices

In modern optical communications systems it is often desirable to combine active and passive functions on the same substrate. Fortunately, many polymer materials exhibit the electro-optic effect, in which the index of refraction is changed by the presence of an electric field. This makes them useful in a variety of electrically controlled switches and modulators. Electro-optic modulators will be discussed in detail in Chapter 9, but it is worthwhile to note at least one of them fabricated in a polymer in this chapter. Morand et al. [25] have made a phase modulator and switch in a hybrid structure that consists of a thin film of electro-optic polymer spin-coated on a passive waveguide made by ion exchange on a glass substrate. In this device the lightwaves actually travel mostly in the ion-exchange glass waveguide, which has dimensions that limit it to single-mode operation. However, the tail of the optical mode extends into the active polymer and provides coupling between the applied electric field and the optical mode. The active zone consists of 4 layers. A thin conductive oxide layer (< 50 nm) is deposited on top of the glass surface to serve as a ground electrode, and its edges are coated with chrome to reduce series resistance.

Then $0.5\text{ }\mu\text{m}$ of electro-optic polymer is deposited by spin coating, followed by a $2\text{ }\mu\text{m}$ thick layer of epoxy resin to protect the active polymer. Finally, an aluminum layer is evaporated on top of the active zone to serve as the signal electrode. When an external field is applied between the signal and ground electrodes, it changes the index of refraction in the polymer, which in turn causes a phase shift in the light-waves traveling in the waveguide. Hence the device functions as a phase modulator. If the waveguide is patterned into a phase sensitive structure such as a Mach Zehnder interferometer (to be discussed in Chapter 9) the change in phase can be converted to a change in amplitude, yielding an optical intensity modulator or switch. In the case of the particular device described above, the signal voltage required to produce a phase change of π radians (180°), which upon conversion to an amplitude modulator would correspond to 100% switching, was just 10 volts. This is similar to what would be required in made in III–V semiconductors or in lithium niobate.

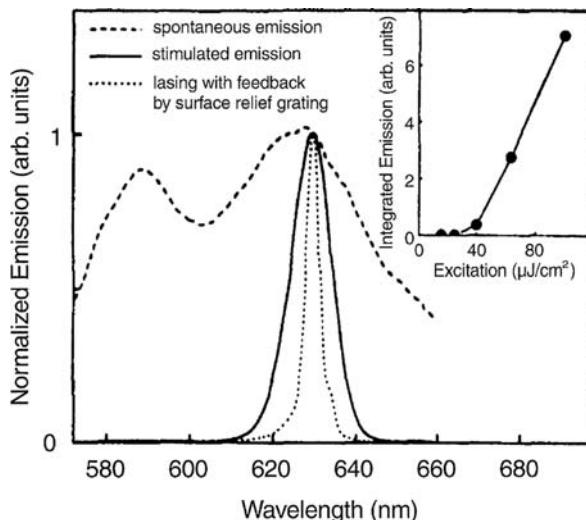
One of the most important active functions that is desired in an optical integrated circuit is that of light emission. At first it was thought that effective light emission required a semiconductor and might not be possible in a polymer. However, recent work has shown that not only incoherent light emitters, but also lasers, can be made in polymers. For example, Shoustikov et al. [26] have reported light emission in a variety of polymers that are capable of emitting in the wavelength range from 462 to 645 nm, and have provided an extensive list of references. They have made organic light emitting diodes (OLEDs) consisting of a hole transporting layer, an emissive layer (doped organic polymer) and an electron transporting layer, sandwiched between a hole injecting ITO (indium tin oxide) contact and an electron injecting contact. These devices exhibit an effective optical band gap similar to that of a semiconductor.

Schulzgen et al. [27] have observed stimulated emission of light in polymers and have made planar and ring lasers of small size that are suitable for applications in integrated optics. (The basic principles of stimulated emission vs. spontaneous emission and lasing that are discussed in detail in Chapters 11 and 12 apply to polymers as well as semiconductors.) Very large optical gain due to stimulated emission was measured directly by pump-probe spectroscopy. Stimulated emission from fewhundred-nanometer thin films resulted in a sharp increase in emission intensity and emission line narrowing, as shown in Fig. 5.10. These are the two standard indications of stimulated emission replacing spontaneous emission. When feedback was provided, further linewidth narrowing and a distinct threshold break in the light output vs. input power curve demonstrated that lasing was occurring.

The excitation source in this case was optical; an amplified colliding pulse mode-locked (CPM) laser system provided tunable pulses of 100 femtosecond duration. The optical gain in the polymer covered a spectral range of 50 nm and was observed to be as high as 10^4 cm^{-1} .

Work on polymer LEDs and lasers has continued at a fast pace in recent years. Semiconducting polymer LEDs [28] are now well established for use as discrete devices [29]. LEDs with CW brightness in excess of $25,000\text{ cd/m}^2$ and a peak pulsed brightness exceeding $2,000,000\text{ cd/m}^2$ have been reported [29]. LEDs are also entering the market place as part of displays [30]. Currently most polymer

Fig. 5.10 Comparison of the emission spectra from a 200-nm BEH:PPV film: Spontaneous emission at low excitation (*dashed line*), stimulated emission without feedback (*solid line*), and laser emission with feedback provided by a surface relief grating (*dotted line*, $\Delta = 500$ nm). *Inset:* Integrated emission intensity of the laser structure as a function of excitation [27]



lasers are optically pumped with high-powered lasers. This approach is useful in the laboratory but is impractical in many applications because of the size and expense of the pump lasers. However, sufficient current densities to achieve threshold have been demonstrated in pulsed applications [31, 32]. For more examples of progress in polymer lasers see, for example, Samuel and Turnbull [33].

5.5 Optical Fiber Waveguide Devices

In addition to being used as interconnections for carrying lightwave signals over both short and long distances, optical fiber waveguides can be used as the basic element of a number of different integrated optic devices. These devices have the inherent advantage that they can be seamlessly coupled to other lengths of optical fiber that are carrying the input and output lightwave signals. For example, Madamopoulos and Riza [34] have made a fiber optic delay line that employs lengths of optical fiber as the delay paths. Lightwave signals are switched through different lengths delay paths as desired. In one configuration, with an electromechanical fiberoptic switch, good optical isolation of 60 dB and low insertion loss of 1.5 dB are obtained for the photonic delay line module, but with a low switching speed of 7 ms. In another configuration, with an electro-optic switch, a faster switching speed of 100 ps is obtained, but optical isolation is reduced to 22 dB and insertion loss is increased to 4 dB.

Fiber Bragg gratings (FBG) are another type of optical fiber device that can perform the functions of reflection, dispersion and deflection of an optical beam. (The Bragg diffraction (reflection) effect is explained in Chapter 15.) A fiber Bragg grating consists of a longitudinal, periodic variation of the refractive index of the core

of an optical fiber waveguide. In a single fiber, the periodicity of the FBG can be chosen to selectively reflect a particular wavelength through 180° while allowing other wavelengths to pass essentially unimpeded. This wavelength selectivity permits them to be used as a component of wavelength division multiplexed systems, even in DWDM systems. Because they have the dispersive characteristics of a grating, FBGs also can be used for dispersion compensation in fiber lightwave systems and for gain flattening in optical amplifiers. In addition, they can be used as very accurate sensors of displacement, mechanical strain, pressure and thermal expansion when they are incorporated into structures such that these factors can cause a change in the periodicity of the grating. Because they are an integral part of the fiber, FBGs have the advantages of simplicity, high coupling efficiency and low cost as compared to similar devices in optical integrated circuit chip form. FBGs can be formed by exposing a step-indexed, germanosilica fiber to intense ultraviolet (UV) light from (typically) a krypton fluoride (KrF) excimer laser ($\lambda = 248 \text{ nm}$) or a frequency-doubled argon-ion laser ($\lambda = 244 \text{ nm}$). Absorption of the light causes a permanent change in the index of refraction of the core. The required accurate periodicity of the FBG can be produced by any of three methods that have been described by Mayer and Basting [35] and are shown in Fig. 5.11. In the interferometric method, a pair of phase coherent laser beams are generated by splitting a single beam in two and using mirrors to fold it back on itself so as to produce an interference pattern of dark and bright bands with the desired spacing. This method has the advantage of being easily tunable to different grating periodicity by changing the angle of the beams, but it is very sensitive to vibration.

The phase-mask method uses a diffraction grating to split a single laser beam into different diffractive orders, which then interfere with each other to produce the grating pattern. This method is less sensitive to vibration than the interferometric method, but a new diffraction grating is required whenever a change in periodicity is desired.

The third method of FBG fabrication utilizes projection of the exposing beam through a patterned mask similar to those used in semiconductor integrated circuit fabrication. The major advantage of this method is that the mask pattern can be shaped so as to produce a “chirped” grating that has multiple periodicities. Such gratings can be used to improve the bandwidth of an optical system and are used in other specialized applications. However, in many applications the periodicity of the grating is such that the spacing between the grating “bars” is on the order of 100 nanometers, and that accuracy is difficult to achieve with a projected mask.

Obviously, each of the three methods has a different set of advantages and disadvantages. An engineer must choose the one that best fits the requirements of a particular application and the capabilities of the equipment that is available.

An example of a specialized application in which a chirped grating is used is the fiber optic accelerometer described by Spammer and Fuhr [36]. The principle of operation is that the chirp width varies as the grating is physically deformed by a displacement x . Sensing is done by passing a relatively broad spectrum light beam through the fiber and monitoring the grating’s peak reflected wavelength in response. Acceleration is given by the second derivative of displacement d^2x/dx^2 ;

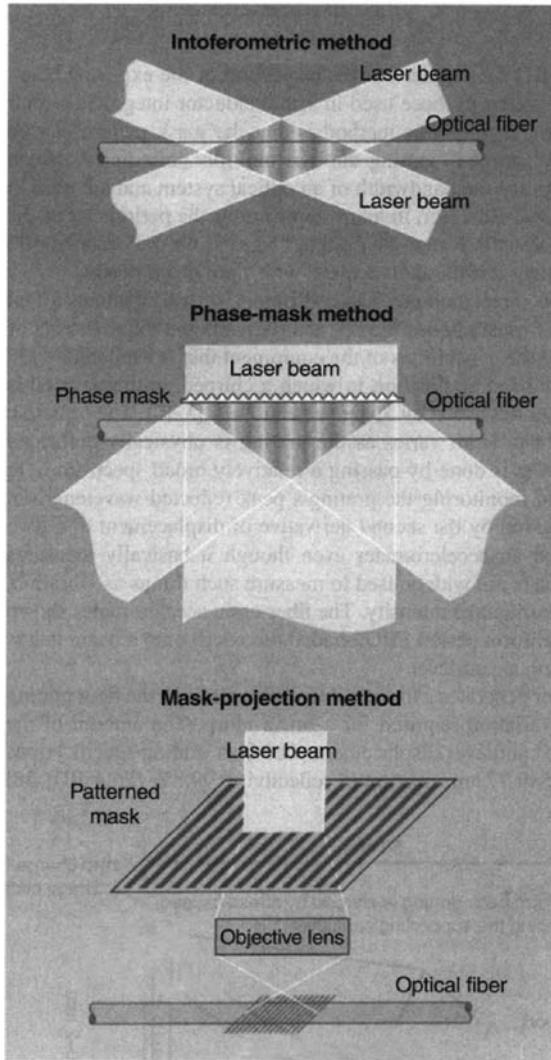


Fig. 5.11 Methods for producing fiber Bragg gratings [35]

hence, the device is called an accelerometer even though it basically measures displacement. Accelerometers are widely used to measure such things as vibration, angles of inclination and earthquake intensity. The fiber optic accelerometer shown in Fig. 5.12 consists of a uniform period FBG bonded (epoxied) onto a beam that is supported at one end to form a cantilever.

Deflecting the cantilever generates a linearly decreasing strain in the fiber grating and results in the period variation required for a linear chirp. The amount of the chirp is proportional to the cantilever displacement. The fiber grating specifications

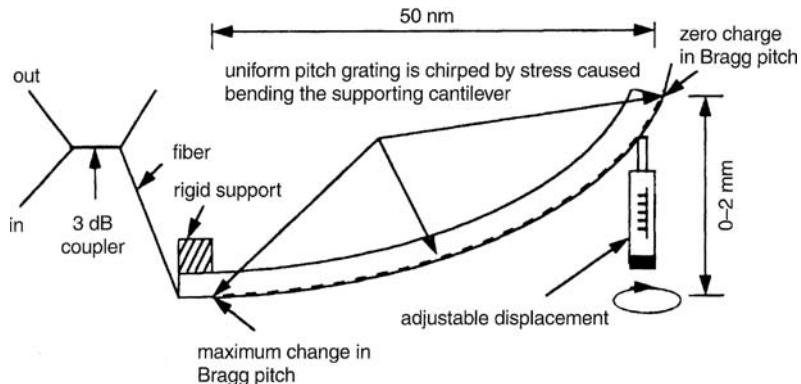


Fig. 5.12 Chirped grating fiber optic accelerometer [36]

are: center wavelength, 1549.77 nm; maximum reflectivity, 99.8% (26.4 dB); full width at half maximum (FWHM), 0.12 nm; and length, 50 mm. The response of this fiber optic accelerometer when exposed to vibration with a fundamental frequency of 20 Hz was compared to that of a conventional piezoresistive accelerometer and found to be essentially the same. The expected linear dependence of peak wavelength on displacement of the cantilever was observed as shown in Fig. 5.13.

When compared to various types of electromechanical accelerometer a fiber optic accelerometer has the advantage that the fiber itself has no moving parts to fail, is very lightweight and is resistant to attack by most chemicals to which it might be exposed.

While glass optical fibers are normally insensitive to chemicals present in their surroundings, dopants can be added to make their optical properties sensitive to a particular chemical. Thus they can be made to serve as chemical sensors. An example of this is the fiber optic chloride sensor of Fuhr et al. [37] which is intended to detect chloride penetration into bridge decks. This is a very important application

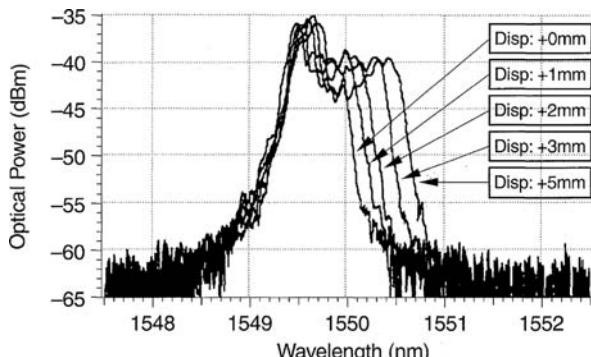


Fig. 5.13 Chirped grating spectra for positive beam tip displacement [36]

to bridge maintenance engineers because the steel strands and bars that are used to reinforce the concrete in bridges are subject to corrosion, and the presence of chlorine ions greatly enhances that corrosion. Since sodium chloride and potassium chloride are often used to melt ice on highways in cold climates, chlorides are likely to be present on bridge surfaces. Glass optical fibers can be made to be sensitive to chlorides by doping them with the organic dye ABQ, a quinoline derivative that exhibits fluorescence that is quenched by the presence of halide ions. The mechanism involved is that of collisional quenching, which is a reversible process whereby the emitted fluorescence intensity decreases nonlinearly with increasing quencher ion concentration. The process is described by the Stern-Volmer relation:

$$I_0/I = 1 + K_{sv}[Q], \quad (5.4)$$

where I_0 is the fluorescence signal in the absence of quencher, I is the fluorescence signal in the presence of quencher concentration $[Q]$, and K_{sv} is the diffusion-dependent Stern-Volmer constant. The nonlinear fluorescence quenching of ABQ with increasing chloride concentration is shown in Fig. 5.14a, and the data are replotted in the Stern-Volmer form of I_0/I in Fig. 5.14b, in which the expected linear curve is observed.

The fluorescence data for ABQ were obtained by stimulating it with a filtered UV source peaked at 340 nm. This might be an inconvenient source to use in the field. However, a similar organic dye, MEQ, can be stimulated to fluorescence by 450 nm radiation, which can be obtained from a LED.

In this chapter, a number of different devices that have been made in either polymers or in glass optical fibers have been reviewed. These are merely a representative sampling of a broad range of devices that fall into this category. It is clear that both active and passive devices can be made, and that often they have advantages in cost and or performance over their counterparts made in the more traditional semiconductor and crystalline dielectric materials. For these reasons one can expect to see

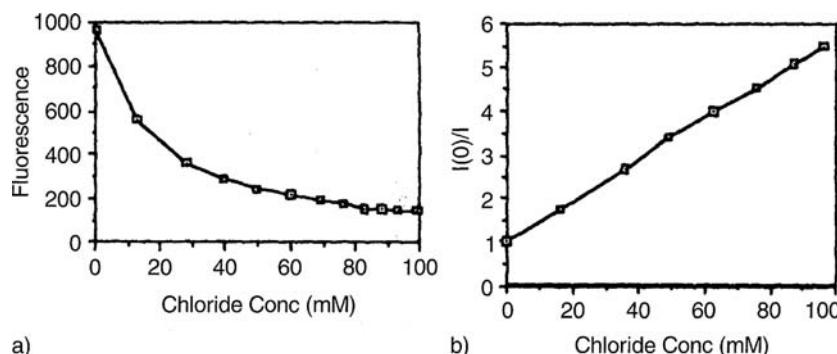


Fig. 5.14 a,b Fluorescence quenching in ABQ by chlorides a nonlinear quenching with increasing chloride concentration b data re-plotted in Stern-Volmer form [37]

even a greater proportion of integrated optic devices to be made in polymers or fibers in the future.

Progress in the design and fabrication of polymer photonic devices has been sufficient to allow them to be incorporated into OICs (or to use the more modern term, PICs, photonic integrated circuits). Bruendel et al. have shown a method of integrating polymer waveguides with polymer light sources [38]. Variable optical attenuator arrays and reconfigurable optical add/drop multiplexer circuits have also been demonstrated in polymers [39]. Polymer waveguides have been used to interconnect arrays of vertical-cavity surface-emitting lasers (VCSELs) [40]. (See Chapter 14 for more information on VCSELs.) Because they are economical and can be patterned effectively, polymers will no doubt be increasingly used in the fabrication of commercial PICs.

Problems

- 5.1 What are the factors that determine the thickness of a polystyrene waveguide formed by spinning a solution of dissolved polystyrene onto a substrate?
- 5.2 An array waveguide multiplexer like that shown in Fig. 5.9 is designed to operate around a center wavelength of $1.55 \mu\text{m}$ and to have 20 channels spaced 1 nm apart in wavelength. The radius of curvature in the focusing slabs is 8 mm . The input and output waveguide spacing is $125 \mu\text{m}$ and the array waveguide separation is $7 \mu\text{m}$. The index of refraction in the slab regions is 1.67. What is the required path length difference between adjacent waveguides in the array?
- 5.3 What can be done in order to double the number of channels in the AWG multiplexer of Prob. 5.2. Describe all of the potential ways that it can be done. Assume that the center wavelength and the wavelength spacing between channels can not be changed.
- 5.4 What advantages do polymers have as compared to III-V semiconductors for use in the fabrication of photonic integrated circuits?
- 5.5 Why is it possible to fabricate LEDs emitting at significantly different wavelengths (i.e., different colors) all in the same polymer?
- 5.6 How can you tell if a light-emitting polymer is lasing?
- 5.7 What is a Fiber Bragg Grating (FBG)? Describe three different ways of producing Fiber Bragg Gratings.
- 5.8 A waveguiding film is deposited on a glass substrate by spinning a 280,000 molecular weight polystyrene with a concentration of 10 ml of solvent / gram of polystyrene at a speed of 3000 rpm for 40 s . The film is then annealed at 110°C for 45 minutes . How thick is the film?
- 5.9 A chloride sensor is made by coating a glass optical fiber with the organic dye ABQ. The fluorescence response of this sensor to chloride concentration is shown in Fig. 5.14. What is the Stern-Volmer constant for this material as determined from the measured data?

References

1. N. Keil, H.H. Yao, H.H.C. Zawadzki, J. Bauer, M. Bauer, C. Dreyer, J. Schneider: Athermal all-polymer arrayed-waveguide grating multiplexer. *Electron. Lett.* **37**, 579 (2001)
2. L. Eldada, L.W. Shaklette: Advances in polymer integrated optics. *IEEE J. Selected Topics in Quant. Electron.* **6**, 54 (2000).
3. J. Everhart: Analysis of polystyrene and polystyrene-poly(methyl methacrylate) diblock copolymer for the creation of optical waveguides. Masters Thesis, University of Delaware (University Microfilms, Ann Arbor, MI 1998)
4. D.P. Prakash, D.C. Scott, H.R. Fetterman, M. Matloubian, Q. Du, W. Wang: integration of polyimide waveguides with traveling-wave phototransistors, *Phot. Technol. Lett.* **9**, 800 (1997)
5. J. Kobayashi, T. Matsuura, Y. Hida, S. Sasaki, T. Maruno: Fluorinated polyimide waveguides with low polarization-dependent loss and their applications to thermooptic switches, *J. Light-wave Technol.* **16**, 1024 (1998)
6. Y-T. Lu, Z-L. Yang, S. Chi: Fabrication of a deep polyimide waveguide grating for wavelength selection, *Opt. Commun.* **216**, 127 (2003)
7. F. Zhang: Low loss coupling between lasers and other optoelectronic devices Masters Thesis. University of Delaware (University Microfilms, Ann Arbor, MI 1999)
8. R.K. Watts: Lithography, in *VLSI Technology*, S.M. Sze (ed.), 2nd edn. (McGraw Hill, New York 1988)
9. R.R. Krchnavek, G.R. Lalk, D.H. Hartman: *J. Appl. Phys.* **66**, 5156 (1989)
10. D.H. Hartman, G.R. Lalk, J.W. Howse, R.R. Krchnavek: *Appl. Opt.* **28**, 40 (1989)
11. J.H. Trewella, J. Gelorme, B. Fan, A. Speth, D. Flagello, M. Oprysko: *SPIE Proc.* **1777**, 379 (1989)
12. A. Guha, J. Bristow, C. Sullivan, A. Husain: *Appl. Opt.* **29**, 1077 (1990)
13. R. Selvaraj: *IEEE J. L T-6*, 1034 (1988)
14. A. Neyer, T. Knoche, L. Muller: *Electron. Lett.* **29**, 399 (1993)
15. D.P. Prakash, D.V. Plant, D. Zhang, H.R. Fetterman: *SPIE Proc.* **1774**, 118 (1993)
16. J.E. Thomson, H. Levesque, E. Savov, F. Horowitz, B.L. Booth, J.E. Marchegiano: *Opt. Eng.* **33**, 939 (1994)
17. B.L. Booth: Optical interconnection polymers, in *Polymers for Lightwave and Integrated Optics: Technology and Applications*, L.A. Hornak (ed.) (Dekker, New York 1992)
18. L. Eldada, R. Blomquist, L. Shaklette, M. McFarland: High performance polymetric componetry for telecom and datacom applications. *Opt. Eng.* **39**, 596 (2000)
19. J.T. Gallo, J.L. Hohman, B.P. Ellerbusch, R.J. Furmanak, L.M. Abbott, D.M. Graham, C.A. Schuetz, B.L. Booth: High-density interconnects for 2-dimensional VCSEL arrays suitable for mass scale production. *SPIE ITCOM2001 Conf. on Modeling and Design of Wireless Networks*, Denver, USA, 23–24 August 2001
20. J.E. van der Linden, P.P. Van Daele, P.M. Dobbelaere, M.B. Diemeer: Compact multichannel in-line power meter. *IEEE Photonics. Tech. Lett.* **11**, 263 (1999)
21. T. Pliska, V. Ricci, A.C. Le Duff, M. Canva, P. Raymond, F. Kajzar, G.I. Stegeman: Low loss polymer waveguides fabricated by plasma etching for nonlinear-optical devices operating at telecommunication wavelengths. *Tech. Digest Quantum Electronics and Laser Science Conference* (1999) p. 138
22. A. Chen, V. Chuyanov, S. Garner, W.H. Steier, J. Chen, Y. Ra, S.S.H. Mao, G. Lan, L.R. Dalton: Fast maskless fabrication of electrooptic polymer devices by simultaneous direct laser writing and electric poling of channel waveguides. *Proc. Lasers and Electro-Optics Society Annual Meeting*, 1997. LEOS'97 **2**, 250 (1996)
23. A. Chen, V. Chuyanov, F. Marti-Carrera, S. Garner, W.H. Steiner, J. Chen, S. Sun, L.R. Dalton: Vertically tapered polymer waveguide mode size transformer for improved fiber coupling. *Opt. Eng.* **39**, 1507 (2000)
24. A. Kaneko, T. Goh, H. Yamada, T. Tanaka, I. Ogawa: Design and applications of silica-based planar lightwave circuits. *IEEE J. Selected Topics in Quantum Elect.* **5**, 1227 (1999)

25. A. Morand, S. Tedjini, P. Benesch, D. Bosc, B. Loisel: Proc. Glass electro-optic polymer structure for light modulation and switching. Proc. SPIE **3278**, 63 (1998)
26. A. Shoustikov, Y. You, M.E. Thompson: Electroluminescence color tuning by dye doping inorganic light emitting diodes. IEEE J. Selected topics in Quantum Electron. **4**, 1077 (1998)
27. A. Schulzgen, C. Spiegelberg, S.B. Mendes, P.M. Allemand, Y. Kawabe, M. Kuwata-Gonokami, S. Honkanen, M. Fallahi, B.N. Kippelen, N. Peyghambarian: Light amplification and laser emission in conjugated polymers. Opt. Eng. **37**, 1149 (1998)
28. R.H. Friend, R.W Gymer, A.B. Holmes, J.H. Burroughes, R.N. Marks, C. Taliani, D.D.C. Bradley, D.A. Dos Santos, J.L. Brédas, M. Lögdlund, W.R. Salaneck: Electroluminescence in conjugated polymers, Nature **397**, 121 (1999)
29. R.B. Fletcher, D.G. Lidzey, D.D.C. Bradley, S. Walker, M. Inbasekaran, E.P. Woo: High brightness conjugated polymer LEDs, Synthetic Metals **111–112**, 151 (2000)
30. J. Ouellette: Semiconducting polymers on display, The Industrial Physicist **7**, 22 (2001)
31. M.D. McGehee, A.J. Heeger: Semiconducting (conjugated) polymers as materials for solid-state lasers, Advanced Materials **12**, 1655 (2000)
32. N. Tessler: Lasers based on semiconducting organic materials, Advanced Materials **11**, 363 (1999)
33. D.W. Samuel, G.A. Turnbull: Polymer lasers: recent advances, Materials Today **7**, 28 (2004)
34. N. Madamopoulos, N.A. Riza: All-fiber connectorized compact fiber optic delay-line modules using three-dimensional polarization optics. Opt. Eng. **39**, 2338 (2000)
35. E. Meyer, D. Basting: Eximer-laser advances aid production of fiber gratings. Laser Focus World **36**, 107 (April 2000)
36. S.J. Spammer, P.L. Fuhr: Temperature insensitive fiber optic accelerometer using a chirped Bragg grating. Opt. Eng. **39**, 2177 (2000)
37. P.L. Fuhr, D.R. Huston, B. MacCraith: Embedded fiber optic sensors for bridge deck chloride penetration measurement. Opt. Eng. **37**, 1221 (1998)
38. M. Bruendel, Y. Ichihashi, J. Mohr, M. Punke, D.G. Rabus, M. Worgull, V. Saile: Photonic integrated circuits fabricated by deep UV and hot embossing, Digest of the IEEE LEOS Summer Topical Meetings **23–25**, 105 (2007)
39. E.A. Dobisz, L.A Eldada (eds.): Nanoengineered polymers for photonic integrated circuits, Proc. SPIE, **5931**, 121 (2005)
40. J.T. Gallo, J.L. Hohman, B.P. Ellerbusch, R.J. Furmanak, L.M. Abbott, D.M. Graham, C.Z. Schuetz, B.L. Booth, High-density interconnects for 2-dimensional VCSEL arrays suitable for mass scale production, Proc. SPIE **4532**, 47 (2001)

Chapter 6

Losses in Optical Waveguides

Chapters 2 and 3 have explained cutoff conditions in waveguides and described the various optical modes which can be supported. Following the question as to which modes propagate, the next most important characteristic of a waveguide is the attenuation, or loss, that a light wave experiences as it travels through the guide. This loss is generally attributable to three different mechanisms: Scattering, absorption and radiation. Scattering loss usually predominates in glass or dielectric waveguides, while absorption loss is most important in semiconductors and other crystalline materials. Radiation losses become significant when waveguides are bent through a curve.

Thus far we have used either light waves or their associated rays to represent the optical fields. However, loss mechanisms can often be described more advantageously by using the quantum mechanical description, in which the optical field is viewed as a flux of particle-like quantized units of electromagnetic energy, or photons [1]. Photons can be either scattered, absorbed or radiated as the optical beam progresses through the waveguide, thus reducing the total power transmitted. When photons are absorbed, they are annihilated by giving up their energy to either the atoms or subatomic particles (usually electrons) of the absorbing material. In contrast, when photons are scattered or radiated, they maintain their identity by changing only their direction of travel, and sometimes their energy (as in Raman scattering). Nevertheless, scattered or radiated photons are removed from the optical beam, thus constituting a loss as far as the total transmitted energy is concerned.

6.1 Scattering Losses

There are two types of scattering loss in an optical waveguide: volume scattering and surface scattering. Volume scattering is caused by imperfections, such as voids, contaminant atoms and crystalline defects, within the volume of the waveguide. The loss per unit length due to volume scattering is proportional to the number of imperfections (scattering centers) per unit length. Also, the volume scattering loss depends very strongly on the relative size of the imperfections, as compared to the wavelength of light in the material. In all but the crudest of waveguides, volume

imperfections are so small compared to wavelength, and so few in number, that the volume scattering loss is negligible compared to surface scattering loss.

6.1.1 Surface Scattering Loss

Surface scattering loss can be significant even for relatively smooth surfaces, particularly in the case of higher-order modes, because the propagating waves interact strongly with the surfaces of the waveguide. This effect can be visualized best by considering the ray-optic description of the guided wave shown in Fig. 6.1. A wave traveling in the guide experiences many bounces. In a length L , the number of reflections from each surface is given by

$$N_R = \frac{L}{2t_g \cot \theta_m}. \quad (6.1)$$

As was demonstrated in Problem 2.6, for the case of a Ta_2O_5 waveguide with $t_g = 3 \mu\text{m}$, $n_2 = 2.0$, and $\beta_m = 0.8 kn_2$, light of wavelength $\lambda_0 = 9000 \text{ \AA}$ undergoes 1250 reflections from each surface for each cm traveled. Scattering loss then occurs at each reflection. Since θ_m is larger for the higher-order modes, they experience greater loss because of surface scattering.

To quantitatively describe the magnitude of optical loss, the exponential attenuation coefficient is generally used. In that case, the intensity (power per area) at any point along the length of the waveguide is given by

$$I(z) = I_0 e^{-\alpha z}, \quad (6.2)$$

where I_0 is the initial intensity at $z = 0$. It can be shown (Problem 6.2) that the loss in dB/cm is related to α by

$$\mathcal{L} \left[\frac{\text{dB}}{\text{cm}} \right] = 4.3\alpha[\text{cm}^{-1}]. \quad (6.3)$$

Tien [2] has derived an expression for scattering loss due to surface roughness, based on the Rayleigh criterion. He has shown that loss to be

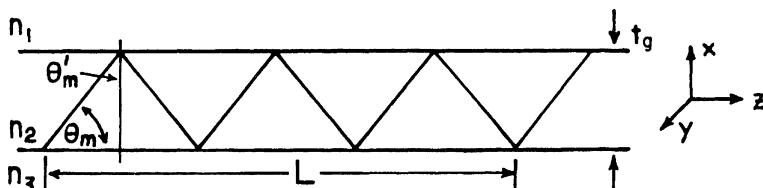


Fig. 6.1 Diagram of a ray optic approach to determination of scattering loss

$$\alpha_s = A^2 \left(\frac{1}{2} \frac{\cos^3 \theta'_m}{\sin \theta'_m} \right) \left(\frac{1}{t_g + (1/p) + (1/q)} \right), \quad (6.4)$$

where θ'_m is as shown in Fig. 6.1, p and q are the extinction coefficients in the confining layers (Section 3.1.1), and A is given by

$$A = \frac{4\pi}{\lambda_2} (\sigma_{12}^2 + \sigma_{23}^2)^{1/2}, \quad (6.5)$$

where λ_2 is the wavelength in the guiding layer, and σ_{12}^2 and σ_{23}^2 , are the variances of surface roughness. It will be recalled that the statistical variance of a variable x is given by

$$\sigma^2 = S[x^2] - S^2[x], \quad (6.6)$$

where $S[x]$ is the mean value of x and

$$S[x^2] = \int_{-\infty}^{\infty} x^2 f(x) dx, \quad (6.7)$$

where $f(x)$ is the probability density function.

The expression given in (6.4) is based on the Rayleigh criterion, which says: if the incident beam at the surface has power P_i , then the specularly reflected beam has power

$$P_r = P_i \exp \left[- \left(\frac{4\pi\sigma}{\lambda_2} \cos \theta'_m \right)^2 \right]. \quad (6.8)$$

This Rayleigh criterion holds only for long correlation lengths, but that is a reasonably good assumption in most cases.

Note in (6.4) and (6.5) that the attenuation coefficient α is basically proportional to the square of the ratio of the roughness to the wavelength in the material, as represented by A^2 . This ratio is then weighted by a factor varying inversely with the waveguide thickness plus the terms $1/p$ and $1/q$, which are related to the penetrating *tails* of the mode. Obviously, well-confined modes may be more affected by surface scattering than modes with large evanescent tails. If $1/p$ and $1/q$ are large compared to t_g , scattering will be reduced. Physically, the penetration of the wave at the interface makes it less sensitive to surface roughness, tending to average out the effect of variations. The factor $\cos^3 \theta'_m / \sin \theta'_m$ in (6.4) accounts for the greater loss for higher-order modes (with small θ'_m) because of more reflections at the surface per unit length traveled in the direction of propagation. Experimentally measured [3] values of α_s for a Ta₂O₅ waveguide, and light of 6328 Å wavelength, correlate very well with those calculated using (6.4), indicating $\alpha_s = 0.3 \text{ cm}^{-1}$ for the $m = 0$ mode, as compared to $\alpha_s = 2.8$ for the $m = 3$ mode.

Although Tien's theoretical model for scattering loss [2] is only an approximation, it provides a convenient closed-form expression for α_s . A more sophisticated theory of surface scattering in slab waveguides has been developed by *Marcuse* as part of a larger effort to calculate total waveguide losses [4–6]. Marcuse's theory treats surface scattering as a form of radiation loss in which irregularities in the surfaces of the waveguide couple energy from propagating modes into radiation modes (and into other propagating modes as well). With proper approximations, and in the limit of long correlation length, the results of Marcuse's theory correlate well with those predicted by (6.4).

The far-field radiation patterns of light scattered by surface irregularities have been studied by Suematsu and Furuya [7], and by Miyanaga et al. [8]. In general, the substrate scattering is highly directional, having many very narrow lobes at specific angles, while the air scattering occurs in a single broad lobe; this air lobe is peaked at an angle with respect to the direction of propagation which is dependent on the correlation length. Gottlieb et al. [9] have experimentally observed the out-of-plane scattering of 7059 glass waveguides on substrates of thermally oxidized silicon, and have found the results to be in agreement with the theoretical predictions of *Suematsu* and *Furuya*.

Surface scattering is generally the dominant loss in dielectric film waveguides, such as glasses and oxides, contributing about 0.5–5 dB/cm to the losses of the lowest-order mode, and more for higher-order modes [10, 11]. This loss is consistent with surface variations of about 0.1 μm , which are typically observed in deposited thin film waveguides. In semiconductor waveguides, thickness variations can usually be held to approximately 0.01 μm , and also, absorption losses are much larger, so that surface scattering is not as important.

6.2 Absorption Losses

Absorption losses in amorphous thin films and in crystalline ferroelectric materials, such as LiTaO_3 or LiNbO_3 , are generally negligibly small compared to scattering loss, unless contaminant atoms are present [12, 13]. However, in semiconductors, significant loss occurs because of both interband or *band edge absorption* and free carrier absorption.

6.2.1 Interband Absorption

Photons with energy greater than the bandgap energy are strongly absorbed in semiconductors by giving up their energy to raise electrons from the valence band to the conduction band. This effect is generally very strong, resulting in absorption coefficients that are larger than 10^4 cm^{-1} in direct bandgap semiconductors. See, for example, the absorption curve for GaAs shown in Fig. 4.8. To avoid interband absorption, one must use a wavelength that is significantly longer than

the absorption edge wavelength of the waveguide material. Chapter 4 discussed the technique of adding additional elements to a binary compound to control the bandgap. Such as adding Al to GaAs to form $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ or adding P and In to form $\text{In}_{(1-x)}\text{Ga}_x\text{As}_{(1-y)}\text{P}_y$. This method is generally effective in III-V and II-VI compounds, as long as care is taken in the proper choice of the additional elements so as to obtain an acceptable lattice match between epitaxial layers. Proper concentrations of the elements must be present to sufficiently shift the absorption edge so that the operating wavelength lies beyond the *tail* of the absorption curve. This effect is illustrated for the case of $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ in Fig. 6.2, where experimentally determined absorption data are presented for wavelengths slightly longer than the absorption edge. A multiple abscissa scale is used in Fig. 6.2 to illustrate the shift of the absorption edge to shorter wavelengths as the Al concentrations is increased. It can be seen, for example, that an Al concentration of $x = 30\%$ results in a reduction of absorption loss to 3 dB/cm for a wavelength of 9000 Å. With no Al, the loss would be about 50 cm^{-1} or 215 dB/cm. The level of absorption loss that can be tolerated, of course, depends on the particular application. However, since most OIC's are sized on the order of a centimeter, 3 dB/cm should be acceptable in most cases. Table 6.1 shows the Al concentration required to reduce interband absorption to less than either 8.6 dB/cm or 3 dB/cm, as desired, for light from three different semiconductor laser sources that have been used in OIC's [17].

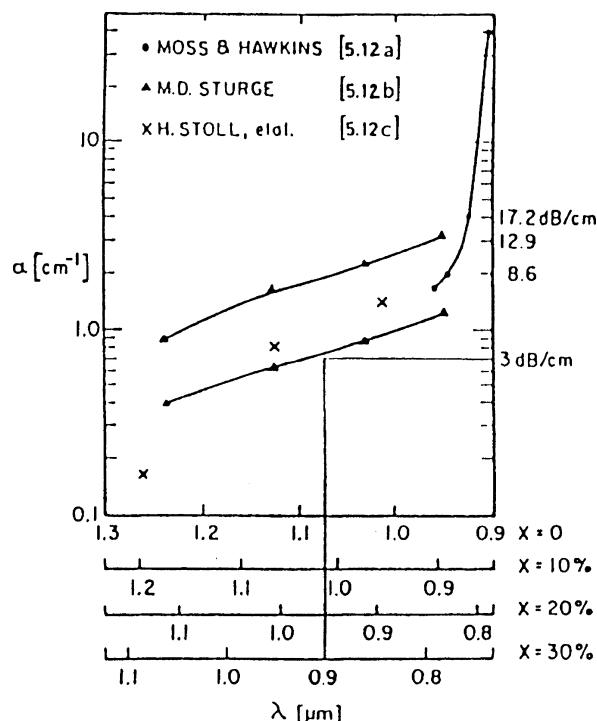


Fig. 6.2 Absorption in the long wavelength tail of the $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ bandedge. Abscissa scales are shown for four different concentrations of aluminum (in atomic percent)

Table 6.1 Absolute aluminum concentration requirements

Source wavelength	Required aluminum concentration in the guide	
	$\alpha = 2 \text{ cm}^{-1}$ (8.6 dB/cm)	$\alpha = 0.7 \text{ cm}^{-1}$ (3 dB/cm)
0.85 μm GaAlAs	17%	40%
0.90 μm GaAs	7%	32%
0.95 - 1.0 μm Si:GaAs	0%	20%

Interband absorption can also be avoided by using a hybrid approach, which employs a laser source emitting a wavelength significantly longer than the absorption edge of the waveguide material. For example, GaAs waveguides have been used for CO₂ laser emission at 10.6 μm [18], and GaP waveguides have been used with He-Ne laser light at 6328 Å [19] and SiC waveguides [20].

Regardless of which approach is used to avoid interband absorption in semiconductor waveguides, additional steps must be taken to eliminate free carrier absorption if practical waveguides are to be realized.

6.2.2 Free Carrier Absorption

Free carrier absorption, sometimes called *intraband* absorption, is that which occurs when a photon gives up its energy to an electron already in the conduction band, or to a hole in the valence band, thus raising it to higher energy. Usually free carrier absorption is taken to include absorption in which electrons are raised out of shallow donor states near the conduction band edge, or holes are excited into the valence band from shallow acceptor states near the valence band edge. The electron transitions between energy states for both interband and intraband (free carrier) absorption are shown in Fig. 6.3.

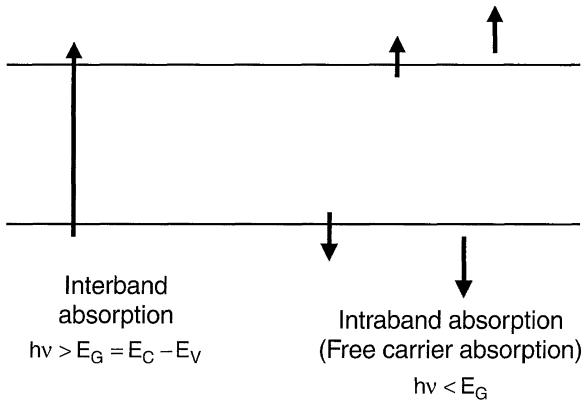
An expression for the absorption coefficient α_{fc} due to free carrier absorption can be derived from classical electromagnetic theory. It is worthwhile to review that derivation because it gives insight into the nature of free carrier absorption and also provides an attendant derivation of the change in index of refraction that is produced by the presence of free carriers.

The motion of an electron in the presence of an applied field $E_0 \exp(i\omega t)$ must satisfy the differential equation [21]

$$m^* \frac{d^2x}{dt^2} + m^* g \frac{dx}{dt} = -e E_0 e^{i\omega t}, \quad (6.9)$$

where g is a damping coefficient and x is the displacement. The first term in (6.9) is the familiar force term (mass \times acceleration); the second term represents a linear

Fig. 6.3 Electron energy transitions during absorption



damping of the electron motion by interaction with the lattice. and the term to the right of the equals sign is the applied force. The steady state solution of (6.9) is

$$x = \frac{(eE_0)/m^*}{\omega^2 - i\omega g} e^{i\omega t}. \quad (6.10)$$

The dielectric constant of a material is given, in general, by

$$K = \frac{\varepsilon}{\varepsilon_0} = 1 + \frac{\bar{P}}{\varepsilon_0 \bar{E}}, \quad (6.11)$$

where \bar{P} is the polarization. In the presence of free carriers

$$\bar{P} = \bar{P}_0 + \bar{P}_1, \quad (6.12)$$

where \bar{P}_0 is the component present without carriers, i.e. the polarization of the dielectric, and \bar{P}_1 is the additional polarization due to the shift of the electron cloud in the field. Thus,

$$K = \frac{\varepsilon}{\varepsilon_0} = 1 + \frac{\bar{P}_0}{\varepsilon_0 \bar{E}} + \frac{\bar{P}_1}{\varepsilon_0 \bar{E}}, \quad (6.13)$$

or

$$K = n_0^2 + \frac{\bar{P}_1}{\varepsilon_0 \bar{E}}, \quad (6.14)$$

where n_0 is the index of refraction of the material without carriers present. Assuming an isotropic material, in which \bar{P} and \bar{E} are in the same direction,

$$\bar{P}_1 = -Ne\bar{x}, \quad (6.15)$$

where N is the free carrier concentration per cm^3 and x is the displacement already given in (6.10). Substituting (6.15) and (6.10) into (6.14), we get

$$K = n_0^2 - \frac{(Ne^2)/(m^*\varepsilon_0)}{\omega^2 - i\omega g}. \quad (6.16)$$

Separating the real and imaginary parts of K , we find

$$K_r = n_0^2 - \frac{(Ne^2)/(m^*\varepsilon_0)}{\omega^2 + g^2}, \quad (6.17)$$

and

$$K_i = \frac{(Ne^2 g)/(m^*\omega\varepsilon_0)}{\omega^2 + g^2}. \quad (6.18)$$

The damping coefficient g can be evaluated from the known steady-state solution of (6.9). Since, at steady state, $d^2x/dt^2 = 0$.

$$m^*g \frac{dx}{dt} = eE. \quad (6.19)$$

From the definition of mobility, μ , we use

$$\frac{dx}{dt} = \mu E. \quad (6.20)$$

which yields

$$g = \frac{e}{\mu m^*}. \quad (6.21)$$

Consider the magnitude of g relative to ω . For the typical case of n-type GaAs, $\mu \cong 2000 \text{ cm}^2/\text{Vs}$ and $m^* = 0.08 m_0$. Thus, $g = 1.09 \times 10^6 \text{ s}^{-1}$. Since $\omega \sim 10^{15} \text{ s}^{-1}$ at optical frequencies, g can most certainly be neglected in the denominators of (6.17) and (6.18). Making that approximation, and also substituting for g from (6.21), we get

$$K_r = n_0^2 - \frac{Ne^2}{m^*\varepsilon_0\omega^2} \quad (6.22)$$

and

$$K_i = \frac{Ne^3}{(m^*)^2\varepsilon_0\omega^3\mu}. \quad (6.23)$$

The exponential loss coefficient α is related to the imaginary part of the dielectric constant by

$$\alpha = \frac{k K_i}{n}, \quad (6.24)$$

where n is the index of refraction and k is the magnitude of the wavevector. Hence, for the case of free carrier absorption,

$$\alpha_{fc} = \frac{k K_i}{n} = \frac{N e^3}{(m^*)^2 n \epsilon_0 \omega^2 \mu c}, \quad (6.25)$$

where $k \equiv \omega/c$ has been used. Since $c = v\lambda_0$ and $\omega = 2\pi v$, (6.25) can be rewritten as

$$\alpha_{fc} = \frac{N e^2 \lambda_0^2}{4\pi^2 n (m^*)^2 \mu \epsilon_0 c^3}. \quad (6.26)$$

For the typical case of $1.15 \mu\text{m}$ light guided in n -type GaAs, with $n = 3.4$, $m^* = 0.08 m_0$, and $\mu = 2000 \text{ cm}^2/\text{Vs}$, (6.26) becomes

$$\alpha_{fc} [\text{cm}^{-1}] \cong 1 \times 10^{-18} N [\text{cm}^{-3}]. \quad (6.27)$$

Thus, free carrier absorption in heavily doped ($N > 10^{18} \text{ cm}^{-3}$)GaAs can be expected to produce losses of the order of $1\text{-}10 \text{ cm}^{-1}$. The major loss due to free carrier absorption occurs in the evanescent tail of optical modes propagating on heavily doped substrates or confining layers (Section 4.3.2). However, by properly choosing the ratio of guide thickness to wavelength one can minimize these losses, as shown by the data of Mentzer et al. [22] given in Fig.6.4.

Before leaving this theoretical development, note in (6.22) that the change in index of refraction resulting from the presence of the carriers is $-(Ne^2)/(m^* \epsilon_0 \omega^2)$, as was given previously, without proof, in (4.3).

The classical expression for α_{fc} given in (6.26) exhibits a λ_0^2 dependence. However, that rarely is exactly observed in practical situations. It must be remembered that the model on which (6.26) is based assumes a constant, wavelength independent, damping coefficient g . In actuality, the damping that occurs because of interaction with the lattice is a varying quantity, which depends on whether acoustic phonons, optical phonons, or ionized impurities are involved. In general, all three will be involved to some extent and the resultant free carrier absorption coefficient can best be represented by [23]

$$\alpha_{fc} = A \lambda_0^{1.5} + B \lambda_0^{2.5} + C \lambda_0^{3.5}, \quad (6.28)$$

where A , B and C are constants giving the relative proportions due to acoustic phonons, optical phonons, and ionized impurities, respectively. Free carrier absorption in n -type compound semiconductors have been studied by Fan [24], and an

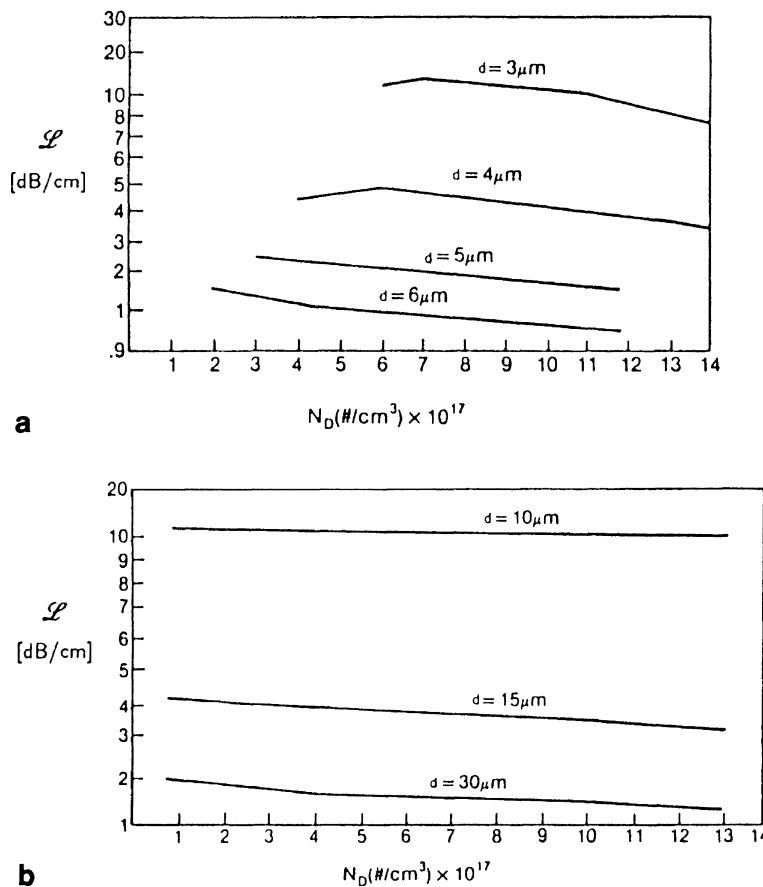


Fig. 6.4 a,b Propagation losses in proton bombarded waveguides (d = waveguide thickness, N_D = substrate doping concentration). **a** $\lambda_0 = 1.3 \mu\text{m}$ propagation losses versus substrate doping. **b** $\lambda_0 = 10.6 \mu\text{m}$ propagation losses versus substrate doping

effective wavelength dependence has been determined for a number of different materials.

In general, (6.26) will give a reasonably accurate estimate of free carrier loss in the visible and near-ir wavelength range, even though it is only an approximation.

6.3 Radiation Losses

Optical energy can be lost from waveguided modes by radiation, in which case photons are emitted into the media surrounding the waveguide and are no longer guided. Radiation can occur from planar waveguides as well as from channel waveguides.

6.3.1 *Radiation Loss from Planar and Straight Channel Waveguides*

Radiation losses from either planar or straight channel waveguides are generally negligible for well confined modes that are far from cutoff. However, at cutoff, all of the energy is transferred to the substrate radiation modes, as discussed in Chapter 2. Since the higher-order modes of a waveguide are always either beyond cutoff or are, at least, closer to cutoff than the lower-order modes, radiation loss is greater for higher-order modes. In an ideal waveguide the modes are orthogonal, so that no energy will be coupled from the lower-order modes to the higher-order modes. However, waveguide irregularities and inhomogeneities can cause mode conversion, so that energy is coupled from lower-order to higher-order modes [25]. In that case, even though a particular mode may be well confined, it may suffer energy loss through coupling to higher-order modes with subsequent radiation. This problem is not usually encountered in typical waveguides of reasonably good quality, and radiation losses can generally be neglected compared to scattering and absorption losses. The one important exception is the case of curved channel waveguides.

6.3.2 *Radiation Loss from Curved Channel Waveguides*

Because of distortions of the optical field that occurs when guided waves travel through a bend in a channel waveguide, radiation loss can be greatly increased. In fact, the minimum allowable radius of curvature of a waveguide is generally limited by radiation losses rather than by fabrication tolerances. Since waveguide bends are a necessary part of all but the simplest OIC's the radiation losses from a curved waveguide must be considered in circuit design.

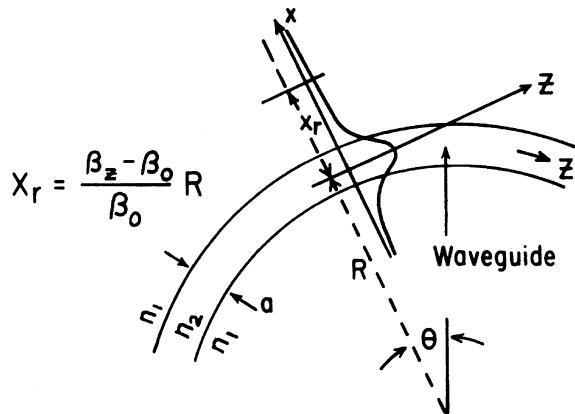
A convenient way of analyzing radiation loss is the velocity approach developed by Marcatili and Miller [26]. The tangential phase velocity of waves in a curved waveguide must be proportional to the distance from the center of curvature because otherwise the phase front would not be preserved. To appreciate this, consider the case of a waveguide mode assumed to be propagating in a circular bend of radius R , with a propagation constant β_z , as shown in Fig. 6.5. There is a certain radius ($R+X_r$) beyond which the phase velocity would have to exceed the velocity of unguided light (in the confining medium, with index n_1), in order to preserve the phase front. Since $d\theta/dt$ must be the same for all waves along the phase front, two resultant equalities are that

$$(R + X_r) \frac{d\theta}{dt} = \frac{\omega}{\beta_0}, \quad (6.29)$$

and

$$R \frac{d\theta}{dt} = \frac{\omega}{\beta_z}, \quad (6.30)$$

Fig. 6.5 Diagram illustrating the velocity approach to the determination of radiation loss



where β_0 is the propagation constant of unguided light in Medium 1, and β_z is the propagation constant in the waveguide at radius R .

Combining (6.29) and (6.30) leads to

$$X_r = \frac{\beta_z - \beta_0}{\beta_0} R. \quad (6.31)$$

The radiation process can be visualized as follows. Photons of the optical mode located at radii greater than $R + X_r$ cannot travel fast enough to keep up with the rest of the mode. As a result, they split away and are radiated into Medium 1. The question which arises is then. "How far must the photons travel before they can be considered as having been removed from the guided mode?". An estimate of this length can be made by analogy to the emission of photons from an abruptly terminated waveguide, as shown in Fig. 6.6. Miller [27] has shown that light emitted into a medium from an abruptly terminated waveguide remains collimated to within a waveguide thickness over a length Z_c given by

$$Z_c = \frac{a}{\varphi} = \frac{a^2}{2\lambda_1}, \quad (6.32)$$

where a and φ are the near field beam width and far-field angle, as shown in Fig. 6.6, and where λ_1 is the wavelength in the medium surrounding the waveguide. The derivation of (6.32) is based on the fundamental relation from diffraction theory

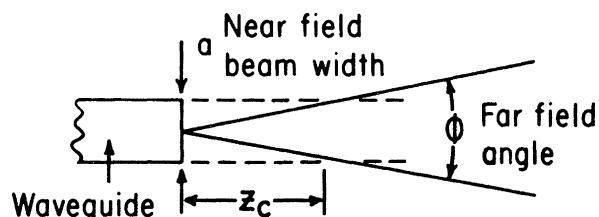


Fig. 6.6 Spread of light emitted from a truncated waveguide

$$\sin \frac{\varphi}{2} = \frac{\lambda_1}{a}, \quad (6.33)$$

which assumes a sinusoidal distribution of fields in the aperture and requires $a > \lambda_1$.

The exponential attenuation coefficient is related to the power lost per unit length traveled in the guide by (see Problem 6.1)

$$\alpha = \frac{-1}{P(z)} \frac{dP(z)}{dz}, \quad (6.34)$$

where $P(z)$ is the power transmitted. Thus, if we define P_1 as the power in the tail of the mode beyond X_r (i.e., the power to be lost by radiation within a length Z_c), and P_t as the total power carried by the waveguide, the attenuation coefficient is given by

$$\alpha \cong \frac{1}{P_t} \frac{P_1}{Z_c}. \quad (6.35)$$

The distance Z_c can be conveniently determined from (6.32), but P_1 must be calculated by integration of the power contained in the optical mode for radii greater than $(R + X_r)$.

If it is assumed that the fields have the form

$$E(x) = \sqrt{C_0} \cos(hx) \quad \text{for } -\frac{a}{2} \leq x \leq \frac{a}{2}, \quad (6.36)$$

and

$$E(x) = \sqrt{C_0} \cos\left(\frac{ha}{2}\right) \exp[-q(|x| - (a/2))] \quad \text{for } |x| \geq \frac{a}{2}. \quad (6.37)$$

then

$$P_1 = \int_{X_r}^{\infty} E^2(x) dx = C_0 \frac{q}{2} \cos^2\left(\frac{ha}{2}\right) \exp\left[-2q\left(X_r - \frac{a}{2}\right)\right] \quad (6.38)$$

and

$$P_t = \int_{-\infty}^{\infty} E^2(x) dx = C_0 \left[\frac{a}{2} + \frac{1}{2h} \sin(ha) + \frac{1}{q} \cos^2\left(\frac{ha}{2}\right) \right]. \quad (6.39)$$

Substituting (6.38) and (6.39) into (6.35) yields

$$\alpha = \frac{\frac{1}{2q} \cos^2\left(\frac{ha}{2}\right) \exp\left(-2q\frac{\beta_r - \beta_0}{\beta_0} R\right) 2\lambda_1 \cdot \exp(aq)}{\left[\frac{a}{2} + \frac{1}{2h} \sin(ha) + \frac{1}{q} \cos^2\left(\frac{ha}{2}\right)\right] a^2}. \quad (6.40)$$

Table 6.2 Waveguide radiation loss data [28, 29]

Case	Index of refraction		Width a [μm]	C_1 [dB/cm]	C_2 [cm $^{-1}$]	R for $\alpha = 0.1$ dB/cm
	Waveguide	surrounding				
1	1.5	1.00	0.198	2.23×10^5	3.47×10^4	4.21 μm
2	1.5	1.485	1.04	9.03×10^3	1.46×10^2	0.78 μm
3	1.5	1.4985	1.18	4.69×10^2	0.814	10.4 cm

While the expression for α in (6.40) appears quite complex, close scrutiny reveals that it has the relatively simple form,

$$\alpha = C_1 \exp(-C_2 R), \quad (6.41)$$

where C_1 and C_2 are constants that depend on the dimensions of the waveguide, and on the shape of the optical mode. The key feature of (6.41) is that the radiation loss coefficient depends exponentially on the radius of curvature. The minimum radius of curvature allowable for radiation loss smaller than 0.1 dB/cm has been calculated by Goell [28] for several typical dielectric waveguides. Experimental data [29] also indicate that such low levels of loss are achievable. However, as can be seen in Table 6.2, the radiation loss can be significant, particularly when the difference in index of refraction between the waveguide and the surrounding medium is very small.

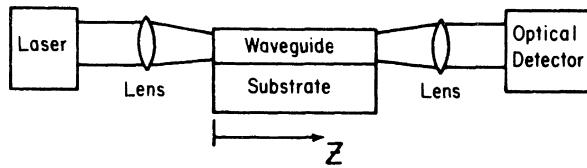
6.4 Measurement of Waveguide Losses

The fundamental method of determining waveguide loss is to introduce a known optical power into one end of the waveguide and measure the power emerging from the other end. However, there are many problems and inaccuracies inherent in using such a basic approach. For example, coupling losses at the input and output are generally not known, and can largely obscure the true waveguide loss per cm. Also, if the waveguide is of the multimode type, losses attributable to individual modes cannot be separately determined. A number of different methods of loss measurement have been devised in order to circumvent problems such as these. The proper choice of a measurement technique depends on what type of waveguide is being used, on what type of loss predominates, and on the magnitude of the loss being measured.

6.4.1 End-Fire Coupling to Waveguides of Different Length

One of the simplest, and also most accurate, methods of measuring waveguide loss is to focus light of the desired wavelength directly onto a polished or cleaved input

Fig. 6.7 Experimental set-up for measurement of waveguide attenuation employing end-fire coupling



face of a waveguide as shown in Fig. 6.7, and then measure the total power transmitted. Such direct coupling is often referred to as *end-fire* coupling. The measurement is repeated for a relatively large number of waveguide samples that have different lengths, but are otherwise identical. Often this series of measurements is performed by beginning with a relatively long waveguide sample, then repetitively shortening the sample by cleaving, or cutting and polishing. Care must be taken before each measurement to align the laser beam and the sample for optimum coupling, by maximizing the observed output power. When this is done, the resulting loss data fall in straight line when plotted on semilog paper, as shown in Fig. 6.8. The loss coefficient can be determined from the slope of the transmission versus length curve, or equivalently, from the relation

$$\alpha = \frac{\ln(P_1/P_2)}{Z_2 - Z_1}, \text{ for } Z_2 > Z_1, \quad (6.42)$$

where P_1 and P_2 are the transmitted power for waveguides of two different lengths Z_1 and Z_2 .

The extent of the scatter of the data points is a measure of the consistency of sample input/output coupling loss, which depends on face preparation, and on sample

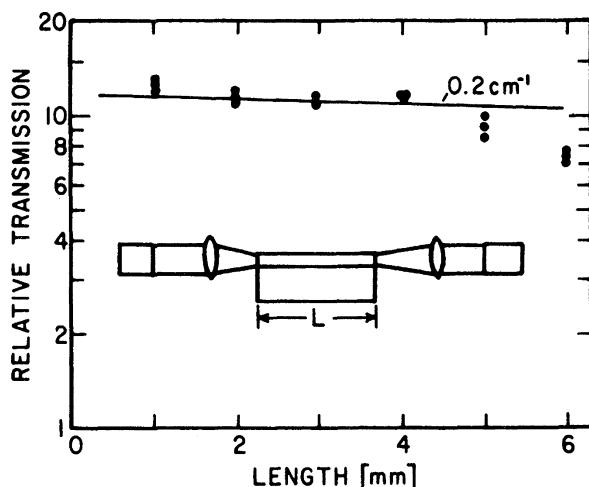


Fig. 6.8 Typical attenuation data. (These data were obtained using the set-up of Fig. 6.7 to measure loss in a Ti diffused waveguide in LiNbO_3)

alignment. If there has been significant change in these parameters from sample to sample, the data points will be widely scattered, and it will be impossible to determine the slope of the transmission versus length curve with acceptable accuracy. However, data points which fall nicely onto a straight line constitute a priori proof that sufficient consistency has been achieved, even though the absolute magnitude of the coupling loss may not be known.

Any departure of the data from the straight line can be taken as an indication of experimental error, assuming that the waveguide is homogeneous. For example, in Fig. 6.8, the increasing differential loss for lengths greater than 4 mm was caused by aperturing of the output beam, which occurred when the light spread laterally in the planar waveguide to an extent that all of it could not be collected by the output lens.

The method of loss measurement described above is most advantageous because of its simplicity and accuracy. However, it does have some inherent disadvantages. Perhaps the most important of these is that it is generally destructive, with the waveguide being chopped to bits in the measurement process. Also, no discrimination is made between the losses in different modes. For these reasons, the technique is most often used for semiconductor waveguides, which are usually single-mode due to the relatively small index difference involved, and which can be easily cleaved to produce end faces of consistent quality.

6.4.2 Prism-Coupled Loss Measurements

In order to determine the loss associated with each mode of a multimode waveguide, the basic measurement technique described in the preceding section can be modified by using prism couplers [30] as shown in Fig. 6.9. Because light can be selectively coupled into each mode by properly choosing the angle of incidence of the laser beam, the loss for each mode can be measured separately. Generally, the position of the input prism is kept fixed while the output prism is moved after each measurement to change the effective sample length. Data are plotted and analyzed just as in the end-fire coupling method. For best accuracy, the output detector should be positioned and masked so as to collect only light from the “*m*-” line corresponding to the desired mode. If there is no mode conversion, all of the transmitted light should be contained in this one *m*-line, and the output prism could be replaced by a

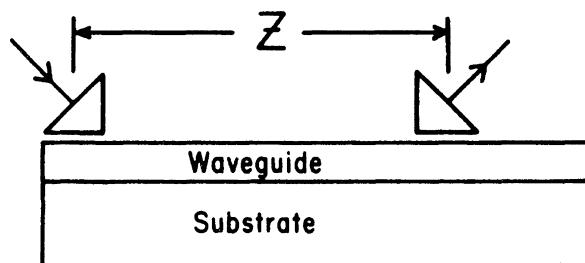


Fig. 6.9 Experimental set-up for measurement of waveguide attenuation employing prism coupling

simpler lens-coupled detector in the end-fire configuration to collect all of the light emerging from the waveguide. In that case, the input prism would have to be moved successively to change the effective sample length. Unless one is certain that mode conversion is insignificant, two prisms should be used. In fact, the use of an output prism is the best way to detect the presence of mode conversion, since multiple m -lines will be observed in that case.

While prism coupling is a more versatile technique for loss measurement than end-fire coupling, it is generally less accurate, because it is very difficult to reproduce the same coupling loss each time the prism is moved to a new position. As will be discussed in more detail in Chapter 7, the coupling efficiency is a strong function of the pressure used to hold the prism in contact with the waveguide; this parameter is difficult to control reproducibly. Nevertheless, the prism coupling method has been used successfully to determine losses as low as 0.02 dB/cm, by employing a small, precisely held prism, with index matching oil between the prism and waveguide [31].

6.4.3 Scattering Loss Measurements

All of the methods of loss measurement that have been discussed thus far determine the *total* loss resulting from scattering, absorption and radiation, with no distinction being made between the three mechanisms. In cases in which absorption and radiation are known to be negligible compared to scattering, such as in straight glass waveguides well above cutoff, it is possible to measure scattering loss conveniently. Since scattered light leaves the waveguide in directions other than the original direction of propagation in the guide, it is possible to determine scattering loss by collecting and measuring the scattered light with an apertured directional detector. One of the most convenient detectors for this purpose is a p-n junction photodiode, coupled to an optical fiber which is used as a probe to collect light scattered from the waveguide [32]. Usually the fiber is held at right angle to the waveguide and scanned along its length so that a plot of relative scattered optical power versus length can be made. The loss per unit length can be determined from the slope of this curve as was done in the case of the previously described methods. This technique of loss measurement implicitly assumes that the scattering centers are uniformly distributed, that the intensity of the scattered light in the transverse direction is proportional to the number of scattering centers, and that radiation and absorption losses are negligible. In that case, it is not necessary to collect all of the scattered light. It is required only that the detector aperture be constant. When an optical fiber is used to probe the waveguide, the spacing between the fiber end-face and the waveguide surface must be kept constant in order to satisfy the requirement of constant detector aperture.

As in the case of loss measurements made with either prism or end-fire coupling, the extent of scatter of the data points about a straight line variation is a reliable indicator of the accuracy of the measurements. If randomly located large scattering centers are present in the waveguide, or if scattering is nonuniform, the loss data

points will not fall neatly into line. The transverse optical fiber probe method is most accurate for waveguides with relatively large scattering loss. When scattering loss is less than about 1 dB/cm, measurements are difficult because of the low intensity that must be detected.

Since scattering loss is usually the dominant mechanism in dielectric thin-film waveguides, this method is used most often to measure losses in such waveguides. In fact, it is usually assumed that the loss measured by the transverse fiber probe method is the total loss in the waveguide. Absorption and radiation loss are assumed to be negligible. In semiconductor waveguides, where absorption loss is more significant, the transverse fiber probe method can be used to determine total loss. However, this will not provide any indication of the relative magnitudes of the three loss mechanisms in the waveguide.

As an alternative to using a transversely oriented optical fiber to collect the light as described in the preceding paragraph, one can use a lens for collecting the scattered light in attenuation measurements. Hurtado-Ramos et al. [33] used an experimental setup of that type to measure losses in SiO_2 and NdF_3 vapor-deposited slab waveguides of $2.7 \mu\text{m}$ thickness, as shown in Fig. 6.10. A prism was used to couple the light from a 20 mW He-Ne laser into the waveguide.

They were able to excite 4 modes in the NdF_3 and 2 in the SiO_2 , but they made all of their attenuation measurements on the lowest order mode, which provided the strongest signal. Losses in the SiO_2 waveguides were measured to be approximately 20 dB/cm , while in the NbF_3 waveguides they were $5\text{-}10 \text{ dB/cm}$.

Continuing research has been directed toward improving loss measurement accuracy and convenience. Chen et al. [34] have used a precision reflectometer to separate the coupling loss between the input fiber and the waveguide from the waveguide propagation loss. Using this technique they measured the propagation loss of a silicon-on-insulator (SOI) rib waveguide fabricated by RIE to be 4.3 dB/cm . A separation of coupling and propagation losses was achieved by Haruna et al. [35]

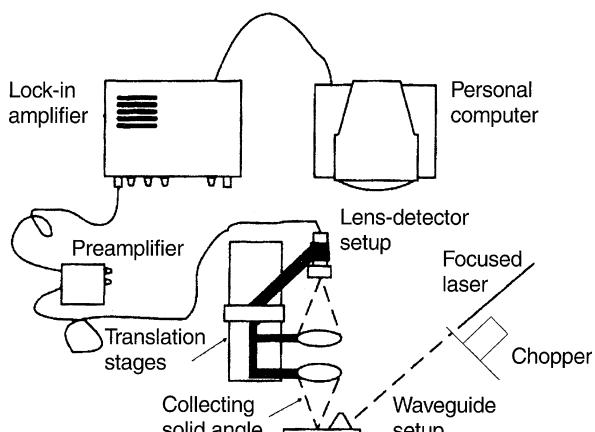


Fig. 6.10 Experimental setup for attenuation measurements [33]

by optimizing the input beam spot size to exactly the same size as the mode size of a guided wave. Using that method, they report achieving accuracy of better than 0.05 dB/cm. Hribek et al. [36] have demonstrated a single-pass transmission loss measurement method in which beam coupling is based on a self-pumped phase conjugation process in a single BaTiO₃ crystal. They report accuracy of at least 10⁻³ cm⁻¹.

In order to explain the various techniques of waveguide loss measurement, it has been necessary to briefly describe some of the methods of coupling light into and out of waveguides. In the next chapter, these methods are described in greater detail, and some additional coupling techniques, such as grating coupling, are reviewed.

Problems

- 6.1 If $P = P_0 \exp(-\alpha z)$, where P_0 is the power at the input end of a waveguide and P is power as a function of distance traveled in the propagation direction (z), show that

$$\alpha = \frac{\text{power lost per unit length}}{\text{power transmitted}}.$$

- 6.2 Show that the relationship between attenuation coefficient α (in cm⁻¹) and loss \mathcal{L} (in dB/cm) is given by

$$\mathcal{L} = 4.3\alpha.$$

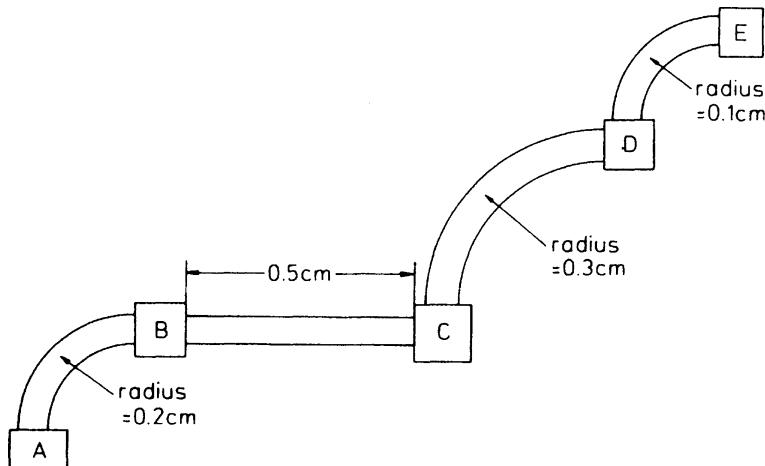


Fig. 6.11 Optical integrated circuit. The curved waveguides are each exactly 1/4 of a circle long for the radii given

- 6.3 In the optical integrated circuit shown above, all of the waveguides have the same cross sectional dimensions and loss per unit length due to scattering and absorption. However, the curved waveguides have an additional loss per unit length due to radiation.

If the total loss between the following elements is:

$$\text{Between } D \text{ and } E \quad L_T = 1.01 \text{ dB}$$

$$\text{Between } C \text{ and } D \quad L_T = 1.22 \text{ dB}$$

$$\text{Between } B \text{ and } C \quad L_T = 1.00 \text{ dB.}$$

What is the total loss L_T between elements *A* and *B*? (Neglect coupling losses - consider only waveguide loss as above.)

- 6.4 Describe the physical reason why radiation loss from a guided optical mode in a curved waveguide increases as the radius of curvature is reduced.
- 6.5 A certain ribbed channel waveguide, 1- μm deep, is used in an OIC for guiding light of vacuum wavelength $\lambda_0 = 6328 \text{ \AA}$. Loss measurements made on test sections of the guide have shown that the loss coefficient in a straight sample is $\alpha = 0.3 \text{ cm}^{-1}$, while in a curved section with radius of curvature $R = 0.5 \text{ mm}$ it is $\alpha = 1.4 \text{ cm}^{-1}$, and in a curved section with $R = 0.3 \text{ mm}$ it is a $\alpha = 26.3 \text{ cm}^{-1}$. What is the minimum radius of curvature that can be used if α must be less than 3 cm^{-1} at all points in the circuit?
- 6.6 (a) Sketch the index of refraction profile in a strip-loaded waveguide, indicating the relative magnitudes of the index in different regions.
 (b) III-V semiconductor waveguide research has progressed to different materials in the order shown in the list below. What reason (i.e., advantage) motivated each step in this progression?
- LPE GaAs waveguides on GaAs substrates
 LPE $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ waveguides on GaAs substrates
 LPE $\text{Ga}_{(1-x)}\text{In}_x\text{As}$ waveguides on GaAs substrates
 MBE $\text{Ga}_{(1-x)}\text{In}_x\text{As}_y\text{P}_{(1-y)}$ waveguides on GaAs substrates
 (LPE - Liquid phase epitaxial; MBE - Molecular beam epitaxial)
- 6.7 A certain uniform waveguide 2-cm long is transmitting an optical signal, the power of which is measured at the output of the waveguide to be 1 W. If the waveguide is cut so that its length is reduced by 10%, the optical power at the output is found to be 1.2 W. What is the attenuation coefficient (in cm^{-1}) in the waveguide?
- 6.8 Describe two different methods for separating the coupling losses from the propagation losses when experimentally measuring the losses of an optical waveguide.
- 6.9 A planar optical waveguide of length = 3 cm is observed to have a transmitted optical power of 5 mW. When the waveguide is cut in half, the optical power at the output of the remaining half is measured to be 5.5 mW.
- (a) What is the loss coefficient of the waveguide in cm^{-1} ?
 (b) What is the loss of the waveguide in dB/cm ?

- 6.10 A planar optical waveguide of length = 5 cm is fabricated in an epitaxial layer of GaAs that has an n-type doping concentration of $1 \times 10^{18}/\text{cm}^3$.
- What is ratio to the input power of the total power lost over the length of the waveguide that results from free-carrier absorption?
 - What is the answer to part (a) if the doping concentration is reduced to $1 \times 10^{16}/\text{cm}^3$?
- 6.11 In a waveguide in which scattering loss due to surface roughness is the dominant loss, is the loss greater for a well-confined mode with a small evanescent tail or for one with an evanescent tail extending further into the confining layer?

References

- R.H. Good, Jr.: *The Encyclopedia of Physics*, R.M. Besancon, (Ed.), 3rd edn. (Van Nostrand Reinhold, New York, 1985) p. 921
- P.K. Tien: Appl. Opt. **10**, 2395 (1971)
- D.H. Hensler, J.D. Cuthbert, R.J. Martin, P.K. Tien: Appl. Opt. **10**, 1037 (1971)
- D. Marcuse: Bell Syst. Techn. J. **48**, 3187, 3233 (1969)
- D. Marcuse: Bell Syst. Techn. J. **49**, 273 (1970)
- D. Marcuse: Bell Syst. Techn. J. **51**, 429 (1972)
- Y. Suematsu, K. Furuya: Electron. Commun. Jpn. **56-C**, 62 (1973)
- S. Miyanaga, M. Imai, T. Asakura: IEEE J. **QE-14**, 30 (1978)
- M. Gottlieb, G. Brandt, J. Conroy: IEEE Trans. **CAS-26**, 1029 (1979)
- D.G. Hall, G.H. Ames, R.W. Modavis: J. Opt. Soc. Am. **72**, 1821 (1982)
- D.D. North: IEEE J. **QE-15**, 17 (1979)
- N.K. Uzunoglu, J.G. Fikioris: J. Opt. Soc. Am. **72**, 628 (1982)
- F. Lu, G. Fu, C. Jia, K. Wang, H. Ma, D. Shen: Lithium niobate channel waveguide at optical communication wavelength formed by multienergy implantation, Opt. Express **13**, 9143 (2005)
- T. Moss, G. Hawkins: Infrared Phys. **1**, 111 (1961)
- M.D. Sturge: Phys. Rev. **127**, 768 (1962)
- H. Stoll, A. Yariv, R.G. Hunsperger, E. Garmire: Proton-implanted waveguides and integrated optical detectors in GaAs. OSA Topical Meeting on Integrated Optics, New Orleans, LA (1974)
- V. Evtuhov, A. Yariv: IEEE Trans. **MTT-23**, 44 (1975)
- R.T. Brown: 0-1 GHz waveguide 10.6 μm GaAs electrooptic modulator, IEEE J. Quant. Electron. **28**, 1349 (1992)
- M. Barnoski, R.G. Hunsperger, R. Wilson, G. Tangonan: J. Appl. Phys. **44**, 1925 (1973)
- G. Pandraud, H.T.M. Pham, P.J. French, P.M. Sarro: PECVD SiC optical waveguide loss and mode characteristics, Opt. Laser Technol. **39**, 532 (2007)
- A. Yariv: *Optical Electronics*, 4th edn. (Holt, Rinehart and Winston, New York, 1991) p. 161
- M.A. Mentzer, R.G. Hunsperger, S. Sriram, J. Bartko, M.S. Wlodawski, J.M. Zavada, H.A. Jenkenson: Opt. Eng. **24**, 225 (1985)
- J.I. Pankove: *Optical Processes in Semiconductors* (Prentice-Hall, Englewood Cliffs, NJ 1971) p. 75
- H.Y. Fan: Effects of free carriers on the optical properties. *Semiconductors and Semimetals* **3**, 409 (Academic, New York 1967)
- D. Marcuse: Bell Syst. Tech. J. **48**, 3187 (1969)
- E.A.J. Marcatili, S.E. Miller: Bell Syst. Tech. J. **48**, 2161 (1969)

27. S.E. Miller: Bell Syst. Tech. J. **43**, 1727 (1964)
28. J.E. Goell: Loss mechanisms in dielectric waveguides, in *Introduction to Integrated Optics*, M.K. Barnoski (ed.) (Plenum, New York 1974) p. 118
29. E. Neumann, W. Richter: Appl. Opt. **22**, 1016 (1983)
30. P.K. Tien, R. Ulrich, R.J. Martin: Appl. Phys. Lett. **14**, 291 (1969)
31. H.P. Weber, F.A. Dunn, W.N. Leibolt: Appl. Opt. **12**, 755 (1973)
32. H. Osterberg, L.W. Smith: J. Opt. Soc. Am. **54**, 1078 (1964)
33. J.B. Hurtado-Ramos, O.N. Stavroudis, H. Wang, G. Gomez-Rosas: Scattering loss measurements of evaporated slab waveguides of SiO₂ and NdF₃ using a prism coupler and angle-limited integrated scattering. Opt. Eng. **39**, 558 (2000)
34. S. Chen, Q. Yan, Q. Xu, Z. Fan, J. Liu: Optical waveguide propagation loss measurement using multiple reflections method, Opt. Commun. **256**, 68 (2005)
35. M. Haruna; Y. Segawa, H. Nishihara: Nondestructive and simple method of optical-waveguide loss measurement with optimization of end-fire coupling, Electron. Lett. **28**, 1612 (1992)
36. P. Hribek, M. Slunecko, J. Schröfel: Planar and channel optical waveguide loss measurement using optical phase conjugation in BaTiO₃, Fiber and Integrated Optics **21**, 323 (2002)

Chapter 7

Waveguide Input and Output Couplers

Some of the methods of coupling optical energy into or out of a waveguide were mentioned briefly in Chapter 6. In this chapter, we shall consider in more detail the various coupling techniques that can be used. The methods that are employed for coupling an optical beam between two waveguides are different from those used for coupling an optical beam in free space to a waveguide. Also, some couplers selectively couple energy to a given waveguide mode, while others are multimode. Each type of coupler has its attendant set of advantages and disadvantages; none is clearly best for all applications. Hence, a knowledge of coupler characteristics is necessary for the OIC user, as well as for the designer.

Coupler fabrication is generally accomplished by using techniques such as photoresist masking, thin film deposition and epitaxial growth, which have been described in Chapter 4; thus, a separate chapter devoted to coupler fabrication is not required. However, certain specialized methods, such as holographic exposure of photoresist to make grating couplers, are discussed in this chapter.

7.1 Fundamentals of Optical Coupling

The principal characteristics of any coupler are its efficiency and its mode selectivity. Coupling efficiency is usually given as the fraction of total power in the optical beam, which is coupled into (or out of) the waveguide. Alternatively, it may be specified in terms of a coupling loss in dB. For a mode-selective coupler, efficiency can be determined independently for each mode, while multimode couplers are usually described by an overall efficiency. However, in some cases it is possible to determine the relative efficiencies for the various modes of a multimode coupler. Thus, the basic definition of coupling efficiency is given by

$$\eta_{cm} \equiv \frac{\text{power coupled into (out of) the } m\text{th order mode}}{\text{total power in optical beam prior to coupling}} \quad (7.1)$$

and coupling loss (in dB) is defined as

$$\mathcal{L}_{cm} \equiv 10 \log \frac{\text{total power in optical beam prior to coupling}}{\text{power coupled into (out of) the } m\text{th order mode}}. \quad (7.2)$$

If the power in each mode cannot be separately determined, overall values of η_{cm} and \mathcal{L}_{cm} are used.

Coupling efficiency depends most strongly on the degree of matching between the field of the optical beam and that of the waveguided mode. This principle can be best illustrated by considering the case of the transverse coupler.

7.2 Transverse Couplers

Transverse couplers are those in which the beam is focused directly onto an exposed cross-section of the waveguide. In the case of a free space (air) beam, this may be accomplished by means of a lens. Transverse coupling of two solid waveguides may be done by butting polished or cleaved cross-sectional faces together.

7.2.1 Direct Focusing

The simplest method of transverse coupling of a laser beam to a waveguide is the direct focusing or *end-fire* approach shown in Fig. 7.1. The waveguide may be of either the planar or channel type, but we assume a planar waveguide for the moment. The transfer of beam energy to a given waveguide mode is accomplished by matching the beam-field to the waveguide mode field. The coupling efficiency can be calculated from the overlap integral [1] of the field pattern of the incident beam and the waveguide mode, given by

$$\eta_{cm} = \frac{\left[\int A(x)B_m^*(x)dx \right]^2}{\int A(x)A^*(x)dx \int B_m(x)B_m^*(x)dx}, \quad (7.3)$$

where $A(x)$ is the amplitude distribution of the input laser beam, and $B_m(x)$ is the amplitude distribution of the m th mode.

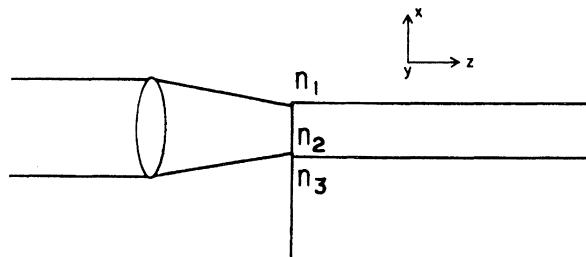


Fig. 7.1 The transverse coupling method, which is sometimes referred to as *end-fire* coupling

The end-fire method is particularly useful for coupling gas-laser beams to the fundamental waveguide mode because of the relatively good match between the Gaussian beam profile and the TE_0 waveguide mode shape. Of course, the beam diameter must be closely matched to the waveguide thickness for optimum coupling. In principle, coupling efficiency could be nearly 100% if field contours were carefully matched. However, in practice, efficiencies of about 60% are usually achieved, because film thicknesses are on the order of $1\text{ }\mu\text{m}$, and thus alignment is very critical. End-fire coupling is often used in the laboratory because of its convenience. However, the difficulty of maintaining alignment without an optical bench limits its usefulness in practical applications.

7.2.2 End-Butt Coupling

Transverse coupling does have a practical application in the case of coupling a waveguide to a semiconductor laser, or to another waveguide. A parallel end-butt approach [2] can be used, as shown in Fig. 7.2. Very efficient coupling can be achieved, since the thickness of the waveguide can be made approximately equal to that of the light emitting layer in the laser, and since the field distribution of the fundamental lasing mode is well matched to the TE_0 waveguide mode. The method is especially useful for coupling a laser diode to a planar waveguide, because efficient coupling of an injection laser to a thin film waveguide is difficult to achieve by using either a prism, grating, or tapered film coupler. The reason for this is that the injection laser has a relatively uncollimated emitted beam which diverges at a half-angle of typically $10\text{--}20^\circ$. Prism, grating and tapered film couplers are all very

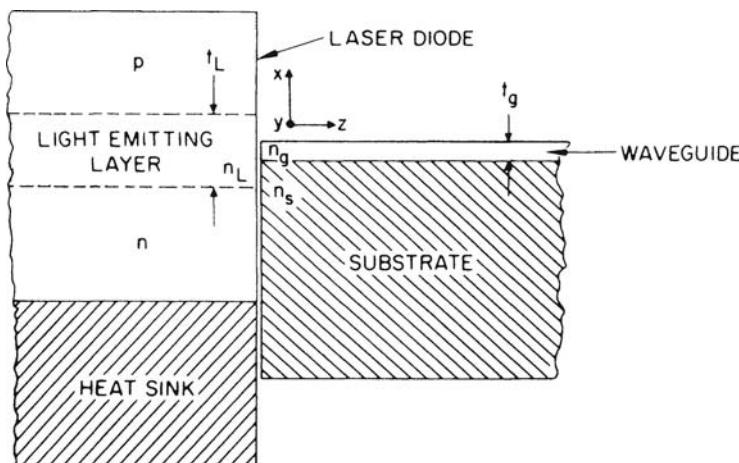


Fig. 7.2 Parallel end-butt coupling of a laser diode and thin-film waveguide

sensitive to the angle of incidence of the light beam, requiring collimation to better than 1° for efficient coupling, as will be explained later in this chapter.

For the case of a laser diode, operating in the fundamental TE_0 mode and coupled to a planar waveguide, as shown in Fig. 7.2, the coupling efficiency for the TE modes is given by [2]

$$\eta_{\text{cm}} = \underbrace{\frac{64}{(m+1)^2 \pi^2}}_{\text{normalization}} \cdot \underbrace{\frac{n_L, n_g}{(n_L + n_g)^2}}_{\text{reflection}} \cdot \underbrace{\cos^2\left(\frac{\pi t_g}{2t_L}\right)}_{\text{overlap}} \cdot \underbrace{\frac{1}{\left[1 - \left(\frac{t_g}{(m+1)t_L}\right)^2\right]^2}}_{\text{overlap}} \cdot \underbrace{\frac{t_g}{t_L}}_{\text{area mismatch}} \cdot \cos^2\left(\frac{m\pi}{2}\right). \quad (7.4)$$

$m = 0, 1, 2, 3, \dots$

The above expression is based on the assumptions that all waveguide modes are well confined, and that $t_g \leq t_L$. It is interesting to note from the last factor of (7.4) that there is no coupling to odd-order waveguiding modes. This is because the field distributions have cancelling lobes when their overlap integrals are taken with the even ($m = 0$) laser mode. The first factor of (7.4) is just a normalization term, while the second factor arises from reflections at the laser-waveguide interface. The other terms account for mismatch in the field distributions in the laser and waveguide.

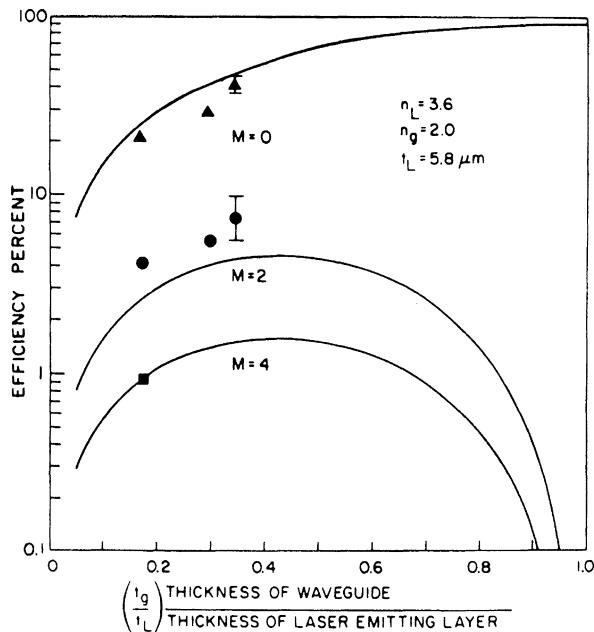
Calculated curves of η_{cm} as a function of relative waveguide thickness are plotted in Fig. 7.3, along with experimental data [2] for the case of a GaAs laser diode coupled to a Ta_2O_5 waveguide on a glass substrate. If $t_g \cong t_L$, coupling efficiency can theoretically approach 100% for the lowest-order waveguide mode. In that case, coupling into higher-order waveguide modes is nearly zero. Hammer et al. have reported using this method to couple as much as 27 mW of optical power from a single-mode diode laser ($\lambda_0 = 0.84 \mu\text{m}$) into a Ti diffused LiNbO_3 waveguide with a coupling efficiency of 68% [3].

The coupling efficiencies shown in Fig. 7.3 are optimum values, corresponding to perfect alignment of the laser and waveguide. Coupling efficiency is most sensitive to transverse lateral misalignment in the x direction. A displacement X of the waveguide relative to the laser, as shown in Fig. 7.4, reduces the coupling efficiency according to the relation [2]

$$P/P_0 = \text{Cos}^2\left(\frac{\pi X}{t_L}\right) \quad (7.5)$$

where P_0 is the coupled power for $X = 0$. The above expression assumes that $t_g < t_L$ and $X \leq (t_L - t_g)/2$. The dashed curve in Fig. 7.4 is the theoretically calculated P/P_0 for the case of $t_L = 5.8 \mu\text{m}$ and $t_g = 2.0 \mu\text{m}$, while the solid curves represent experimentally measured data. A prism output coupler was used on the waveguide

Fig. 7.3 Comparison of experimental coupling efficiency data with theoretical curves as a function of waveguide thickness [2]

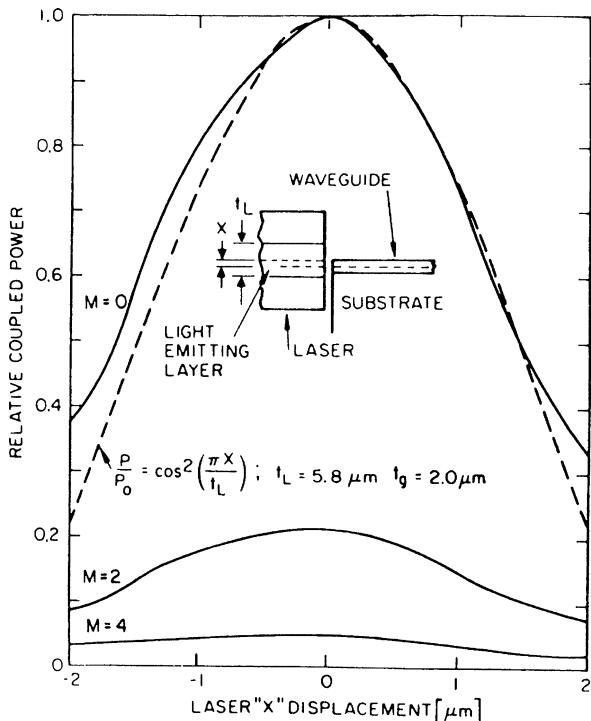


to determine the relative power that was butt coupled into each of the three modes that were observable.

The spacing between the laser and the waveguide in the z direction is also very critical, and must be controlled to a precision on the order of a wavelength for optimum coupling. Figure 7.5 shows the experimentally measured variation of coupled power as a function of z displacement. The oscillatory shape of the curve results from modulation of the effective reflectivity of the laser output face by resonance in the Fabry-Perot etalon formed by the plane parallel faces of the laser and waveguide. In principle, this effect could be eliminated by using an index matching fluid between the laser and waveguide, so that the coupled power would vary smoothly with z displacement, as shown by the dashed curve in Fig. 7.5.

The results presented so far demonstrate that end-butt coupling can be a very efficient means of coupling a diode laser to a waveguide. Enochs [4] has found similar results for the case of butt coupling a laser to an optical fiber. Yet, it is also obvious that submicrometer alignment tolerances are required if optimum efficiency is to be obtained. Alignment to such tolerances can be achieved by using piezoelectrically driven micrometer heads, which feature a small piezoelectric crystal stage bonded onto the end of a conventional screw micrometer. Coarse alignment is established with the micrometer screw, and then final alignment is produced by applying a voltage to the piezoelectric stage to move the laser (or waveguide), while coupled optical power is monitored by means of an appropriate output coupler and photodetector. Piezoelectric micrometers with a sensitivity of better than 40 Å per volt, over

Fig. 7.4 Comparison of experimental coupling efficiency data (solid line) with theoretical curve (dashed) as a function of lateral misalignment of laser and waveguide [2]



a 2000 V range, are commercially available. Thus, alignment to better than 0.1 μm can be permanently bonded to the waveguide support structure with epoxy, or with a metallic bond. Because the size and mass of both the laser and the OIC are relatively small, vibration sensitivity is not a significant problem, and reliable alignment can be maintained.

While the above-described alignment technique is an effective method of coupling a laser diode to a waveguide it is relatively time consuming, and hence expensive in a production line setting. To solve this problem, techniques of both hybrid and monolithic integrated fabrication have been developed which take advantage of automated, batch-fabrication procedures. For example, Yanagisawa et al. [5] have used a hybrid-integration process to couple an AlGaAs laser diode to a glass waveguide on a silicon substrate. Vertical alignment is achieved by using the surface of the silicon substrate as a reference plane, while lateral alignment is done with conventional photolithographic techniques. Alignment accuracy is within 1 μm, leading to a coupling loss of approximately 3 dB.

An example of monolithic integration is the waveguide coupling of an InGaAsP distributed-feedback laser and electro-absorption modulator on an InP substrate described by Aoki et al. [6]. Their method is based on carefully controlled selective-area *Metal-Organic-Vapor-Phase Epitaxy* (MOVPE) of *Multiple-Quantum-Well* (MQW) structures. (See Chapter 18 for a discussion of MQW structures.) The

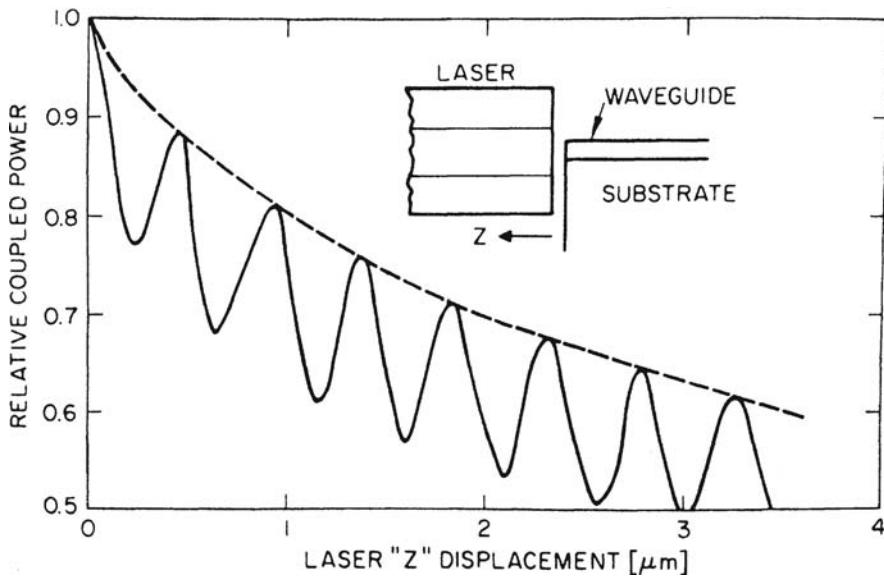


Fig. 7.5 Experimentally measured dependence of coupling efficiency on spacing between laser and waveguide [2]

integrated laser/modulator was employed to transmit data over an 80 km single-mode fiber at a data rate of 2.5 Gbit/s.

7.3 Prism Couplers

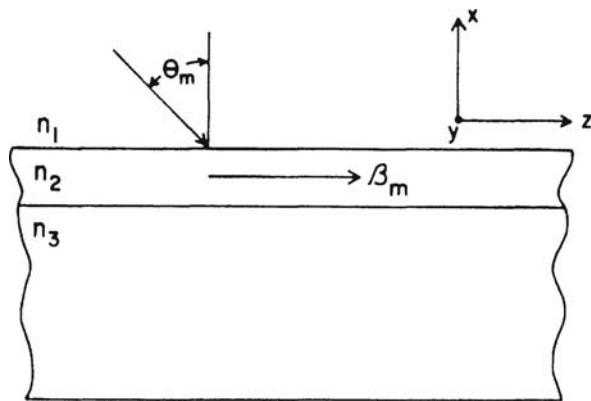
Transverse coupling can be used only when a cross-sectional end face of the waveguide is exposed. In many cases, it is necessary to couple light into a waveguide that is buried within an OIC, with only the surface exposed. One could envision focusing the light onto the surface of the waveguide at an oblique angle, as shown in Fig. 7.6, but a fundamental problem is encountered in that case. For coupling to occur, it is necessary that the components of the phase velocities of the waves in the z direction be the same in both the waveguide and the beam. Thus, a phase-match condition must be satisfied, which requires

$$\beta_m = kn_1 \sin \theta_m = \frac{2\pi}{\lambda_0} n_1 \sin \theta_m. \quad (7.6)$$

However, in Chapter 2 it was shown that, for a waveguided mode,

$$\beta_m > kn_1. \quad (7.7)$$

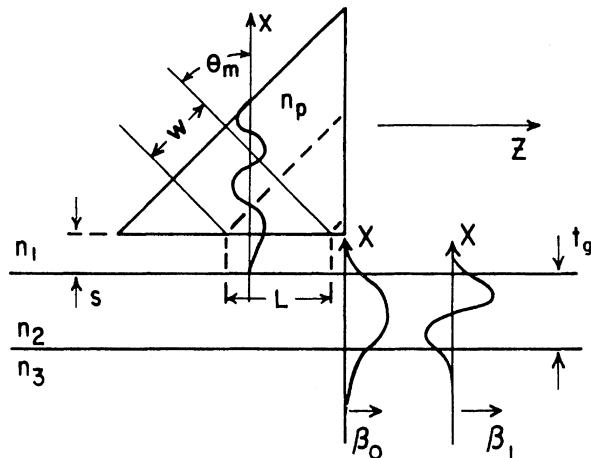
Fig. 7.6 Diagram of an attempt to obliquely couple light into a waveguide through its surface



Combining (7.6) and (7.7) leads to the result that $\sin \theta_m > 1$, which is, of course, impossible.

One solution to the problem of phase matching is to use a prism, as shown in Fig. 7.7. A beam of light of width W is directed into the face of the prism, which has $n_p > n_1$. The beam is totally internally reflected at the $n_p - n_1$ interface, setting up a standing wave mode in the prism, as shown in Fig. 7.7. This mode is stationary in the x direction, but moves in the z direction with a phase constant β_p . In the waveguide, various guided modes can exist, moving in the z direction with phase constants β_m . All of these guided modes have an evanescent tail extending slightly beyond the $n_1 - n_2$ interface. If the prism spacing s is small enough so that the tails of the waveguide modes overlap the tail of the prism mode, there is coherent coupling of energy from the prism mode to the m th waveguide mode when θ_m is chosen so that $\beta_p = \beta_m$. The condition for matching of the β terms is given by

Fig. 7.7 Diagram of a prism coupler. The electric field distributions of the prism mode and the $m = 0$ and $m = 1$ waveguide modes in the x direction are shown



$$\frac{2\pi n_p}{\lambda_0} \sin \theta_m = \beta_m. \quad (7.8)$$

Although θ_m must be carefully chosen in order to couple to a given mode, a single prism can be used to couple to many different modes by merely changing the angle of incidence of the optical beam. The beam need not be perpendicular to the prism surface, as shown in Fig. 7.7. However, if the beam is not perpendicular to the prism surface, refraction at that interface will require a modification of the expression for θ_m given by (7.8). See Problem 7.2 for an example of this effect.

The process of coupling energy via the overlapping mode tails, while the incident beam tends to be totally internally reflected in the prism, is sometimes called *optical tunneling*, because it is analogous to the quantum mechanical tunneling of a particle through an energy barrier. The modes in the waveguide are only weakly coupled to the mode in the prism. Hence, negligible perturbation of the basic mode shapes occurs. Of course, the condition

$$\theta_m > \theta_c = \sin^{-1} \left(\frac{n_1}{n_p} \right) \quad (7.9)$$

must also be satisfied if total internal reflection is to occur in the prism, where θ_c is the critical angle.

Because of the size of the prism, the interaction between prism and waveguide modes can occur only over the length L . The theory of weakly coupled modes [1] indicates that a complete interchange of energy between phase-matched modes occurs if the interaction length in the z direction satisfies the relation

$$\kappa L = \pi/2. \quad (7.10)$$

where κ is the coupling coefficient. The coefficient κ depends on n_p , n_1 and n_2 , which determine the shape of the mode tails, and on the prism spacing s . From (7.10), the length required for complete coupling is given by

$$L = \frac{W}{\cos \theta_m} = \frac{\pi}{2\kappa}. \quad (7.11)$$

For a given L , the coupling coefficient required for complete coupling is thus given by

$$\kappa = \frac{\pi \cos \theta_m}{2 W}. \quad (7.12)$$

This condition for complete coupling assumes that the amplitude of the electric field is uniform over the entire width W of the beam. In a practical case this is never true. For a Gaussian beam shape, it results that the maximum coupling efficiency is about 80%. For a more detailed discussion of the effect of beam width and shape on

coupling efficiency see Tamir [7], and Klimov et al. [8]. Also, it can be seen that, in order to get 100% coupling with a uniform beam, the trailing edge of the beam must exactly intersect the right-angle corner of the prism. If it intersects too far to the right, some of the incident power will be either reflected or transmitted directly into the waveguide and will not enter the prism mode. If the beam is incident too far to the left, some of the power coupled into the waveguide will be coupled back out into the prism.

The prism coupler is frequently used in integrated optics applications because of its versatility. It can be used as either an input or output coupler. When used as an output coupler, the prism is arranged exactly as in Fig. 7.7, except that the direction of travel of the waveguided light would be in the negative z direction. If more than one mode is propagating in the guide, light is coupled out at specific angles corresponding to each mode. Because of this characteristic, the prism coupler can be used as an analytical tool to determine the relative power in each waveguide mode, as described in Chapter 6. The prism can also be moved along the length of the waveguide to determine losses. However, care must be taken to apply the same mechanical pressure to the prism during each measurement of coupled power so that the spacing, and hence the coupling coefficient, will be constant.

One disadvantage of the prism coupler is that n_p must be not only greater than n_1 but also greater than n_2 . This is true because the waveguide index n_2 is generally close to the substrate index n_3 , which leads to the result that

$$\beta_m \cong kn_2 = \frac{2\pi}{\lambda_0} n_2. \quad (7.13)$$

Since $\sin \theta_m \leq 1$, (7.13) coupled with (7.8) implies that $n_p > n_2$. In the case of glass waveguides, with indices $\cong 1.5$, it is easy to find a suitable prism material with $n_p > n_2$. However, semiconductor waveguides, which typically have indices $\cong 3$ or 4, are more difficult to couple with prisms. Both the index and the transparency of the prism material must be considered at the wavelength of interest. Table 7.1 gives a number of different prism materials that are available in good optical quality, along with their indices of refraction at various wavelengths.

Another disadvantage of the prism coupler is that the incident beam must be highly collimated because of the critical angular dependence of coupling efficiency into a given mode. Due to this problem, prism couplers cannot be used effectively

Table 7.1 Practical prism materials for beam couplers

Material	Approximate refractive index	Wavelength range
Strontium titanat	2.3	visible–near IR
Rutile	2.5	visible–near IR
Germaium	4.0	IR

with semiconductor lasers, which have a beam divergence half angle of 10–20° unless a lens is used to collimate the beam.

Prism couplers can be used in some situations for effectively coupling optical beams to other than planar waveguide structures. They have recently been used to couple light into and out of the fundamental whispering-gallery mode (WGM) of a dielectric microsphere [9]. Light from a tunable narrow-linewidth diode laser emitting around 670 nm was focused by a lens to a small spot on the surface of a triangular prism that was in contact with a 59- μm -diameter glass microsphere. This produced an evanescent wave via frustrated total internal reflection within the prism, through which light was coupled to the WGMs of the microsphere resonator.

Prism couplers are very useful in laboratory applications where flexibility is desired regarding the position of the incident beam. However, the requirement of a stable mechanical pressure to hold the prism in place makes it less useful in practical applications, in which vibration and temperature variations are usually encountered. The grating coupler can be used to avoid this problem, without giving up the advantage of mode selectivity.

7.4 Grating Couplers

The grating coupler, like the prism coupler, functions to produce a phase matching between a particular waveguide mode and an unguided optical beam which is incident at an oblique angle to the surface of the waveguide, as shown in Fig. 7.8. It will be recalled that, without the grating, the phase matching condition is given by (7.6), and cannot be satisfied for any θ_m .

7.4.1 Basic Theory of the Grating Coupler

Because of its periodic nature, the grating perturbs the waveguide modes in the region underneath the grating, thus causing each one of them to have a set of spatial harmonics [10] with z -direction propagation constants given by

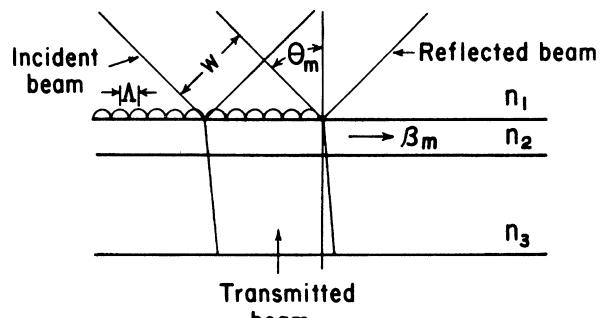


Fig. 7.8 Diagram of a grating coupler

$$\beta_v = \beta_0 + \frac{v2\pi}{\Lambda}, \quad (7.14)$$

where $v = 0, \pm 1, \pm 2, \dots$, and where Λ is the periodicity of the grating. The fundamental factor β_0 is approximately equal to the β_m of the particular mode in the waveguide region not covered by the grating. Because of the negative values of v , the phase matching condition (7.6) can now be satisfied so that

$$\beta_v = kn_1 \sin \theta_m, \quad (7.15)$$

even though $\beta_m > kn_1$.

Since all of the spatial harmonics of each mode are coupled to form the complete surface wave field in the grating region, energy introduced from the beam into any one of the spatial harmonics is eventually coupled into the fundamental ($v = 0$) harmonic as it travels to the right and past the grating region. This fundamental harmonic is very close to, and eventually becomes, the β_m mode outside of the grating region. Thus, the grating coupler can be used to selectively transfer energy from an optical beam to a particular waveguide mode by properly choosing the angle of incidence. The grating can also be used as an output coupler, because, by reciprocity, energy from waveguide modes will be coupled out at specific angles, θ_m , corresponding to a particular mode.

The preceding paragraphs have described the operation of the grating coupler in relatively simple terms. However, the details of the coupling phenomena are quite complex, and depend very strongly on the cross-sectional shape of the grating bars as well as on their spacing. For a detailed discussion of the effect of grating shape, see [7, pp. 110–118]. As in the case of the prism coupler, an optimum coupling efficiency of approximately 80% is theoretically possible when coupling a Gaussian beam with a grating. However, typical unblazed gratings (with symmetric profiles) generally have efficiencies of 10–30%. The principal reason for this is that much of the incident energy is usually transmitted through the guide and lost in the substrate, because, unlike a prism, the grating does not operate in a total internal reflection mode. Power can also be coupled into higher-order diffracted beams produced by the grating, unless the ratio of grating periodicity to guide wavelength is approximately 1.

The efficiency of a grating coupler can be greatly improved by shaping its profile asymmetrically to ‘blaze’ it for optimum performance at the coupling angle and wavelength of interest. For example, Tamir and Peng [11] have shown that the theoretical maximum efficiency for coupling either the TE_0 or TM_0 mode to an air beam is roughly 50% for symmetric grating profiles, while an asymmetric saw-tooth profile can produce an efficiency greater than 95%. These theoretical predictions are supported by experimental results indicating very high efficiencies with blazed gratings [12, 13] and by further theoretical work [14, 16].

The principal advantage of the grating coupler is that, once fabricated, it is an integral part of the waveguide structure. Hence, its coupling efficiency remains constant and is not altered appreciably by vibration or ambient conditions. Also, the

grating coupler can be used on high-index semiconductor waveguides for which it is difficult to obtain a suitable prism material. However, since it is highly angle dependent, the grating coupler cannot be used effectively with the relatively divergent beam of a semiconductor laser. Perhaps the greatest disadvantage of the grating coupler is that it is difficult to fabricate, requiring the use of sophisticated masking and etching techniques.

7.4.2 Grating Fabrication

A grating structure may be formed either by masking and etching the waveguide surface [12,17,18] or by masking the surface and depositing a thin film grating pattern [19, 20]. In either case, the most difficult part of the process is defining the pattern of the closely spaced grating bars. The spacing should be on the order of a wavelength in the waveguide material. Hence, for visible and near ir wavelengths, spacing of 1000–3000 Å are typical, for indices in the range from 1.4 to 4. Conventional photoresists, used in the microelectronics industry, have adequate resolution, but the practical limit of conventional photomasks is approximately 3000 Å. Therefore, gratings are often produced by using either focused e-beam exposure [18] or an optical interference exposure process [19, 20], sometimes called a *holographic* process [21, 22].

In this process the substrate containing the waveguide is first coated with photoresist using any of the methods described in Chapter 4. Then the resist is exposed using an interference pattern generated by combining coherent laser beams directed at the surface as shown in Fig. 7.9. (Both beams are usually obtained from one laser by using a beam splitter.) Simple geometric considerations show that the relationship between grating periodicity A and the beam angle α is given by

$$A = \frac{\lambda_0}{2 \sin \alpha}. \quad (7.16)$$

From (7.16) it is obvious that A is limited to values greater than $\lambda_0/2$. However, by using a rectangular prism, as shown in Fig. 7.10, it is possible to expose grating patterns that are more closely spaced than would ordinarily be possible using a given laser. In this case, (7.16) becomes

$$A = \frac{\lambda_0}{2n \sin \alpha}, \quad (7.17)$$

where n is the index of refraction of the prism material. Using a quartz prism and a He-Cd laser, Yen et al. [23] have produced grating spacings as small as 1150 Å.

Once the photoresist has been exposed, it is developed using standard methods to produce the required mask on the waveguide surface. Either chemical or ion beam etching can be used to produce the grating, following etching procedures described in Chapter 4. In general, ion beam etching produces more uniform grating struc-

Fig. 7.9 Holographic photoresist exposure. The interference pattern produced by two obliquely intersecting coherent laser beams is used to expose the photoresist to produce a grating pattern

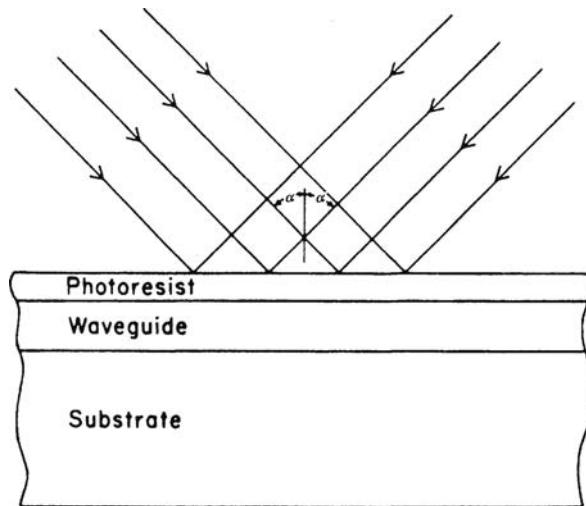
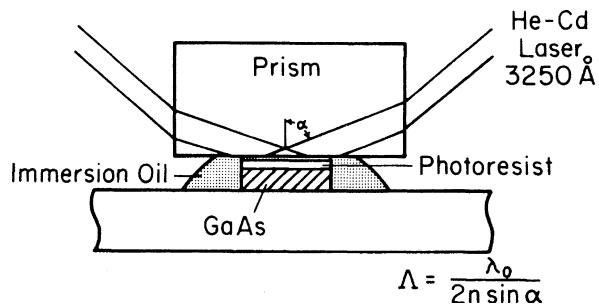


Fig. 7.10 Diagram of prism method of holographic grating exposure as applied to the fabrication of a first order grating in a GaAs waveguide [23]



tures, but chemical etching produces less damage in the waveguide material. Chemical etching is also capable of penetrating more deeply without undercutting, if the etchant is chosen carefully. For example, a buffered solution of NH_4F can be used to etch GaAs to depths greater than 100 μm [24]. Another advantage of chemical etching is that, by proper choice of substrate orientation and etchant, one can obtain non-isotropic etch rates, which yield an asymmetric grating blazed for a desired wavelength [12].

As an alternative to etching the grating structure into the waveguide surface, it can also be produced by depositing thin-film bars on the waveguide surface, using a photoresist mask to define the shape. In fact, the photoresist itself can even be used as a deposited grating material [19, 20]. These methods generally yield a grating coupler with greater scattering and absorption losses than those obtained with an etched grating, because a portion of the beam is blocked from the waveguide.

7.5 Tapered Couplers

The tapered coupler [20] is based on the principle that a waveguide which is below cutoff transfers energy into radiation modes. The waveguide thickness is tapered in the coupler region to produce a reduced-height waveguide with a decreasing cutoff wavelength, as shown in Fig. 7.11. The coupling mechanism can be conveniently visualized by using the ray-optic approach. A guided wave incident on the tapered coupler undergoes zigzag bounces with the angle of incidence to the guide-substrate interface (measured from the normal) steadily decreasing. When the angle of incidence becomes less than the critical angle for total internal reflection, energy is refracted into the substrate. The energy from subsequent rays is also refracted out of the waveguide in a like manner, so that up to 70% coupling efficiency is obtainable [25]. The 30% loss is associated mostly with scattering into air radiation modes when the thickness of the guide reaches the appropriate cutoff point. The waveguide mode is totally coupled out within about 8 vacuum wavelengths after the cutoff point is reached [10].

The greatest advantage of the tapered coupler is that it is simple to fabricate and functions reasonably well as an output coupler. However, it forms a divergent beam, as shown in Fig. 7.11, which spreads over an angle of 1–20°, depending on the taper. This divergent beam is somewhat inconvenient to use, but can be tolerated in many applications in which the form of the output beam is not a critical factor.

In principle, the tapered coupler can also be used as an input coupler. However, in order to obtain high efficiency, one would have to construct a converging input beam which was the reciprocal of the diverging beam shown in Fig. 7.11. Since this is practically an impossibility, only very low efficiencies are usually observed when tapered waveguides are used as input couplers. One practical application of the tapered coupler may be in coupling a thin film waveguide to an optical fiber, since the end of the fiber can be located very close to the waveguide and the end face can be shaped to improve the coupling efficiency.

Tapered couplers can be made relatively easily in optical fiber waveguides by heating the fiber to a temperature at which it softens and then drawing it out in length by applying a tensile force to its ends. This causes the diameter to decrease and the cladding to thin. In this way the cladding can be thinned enough to permit the evanescent tail of the mode(s) in the core to extend outside the fiber and couple to the surrounding media.

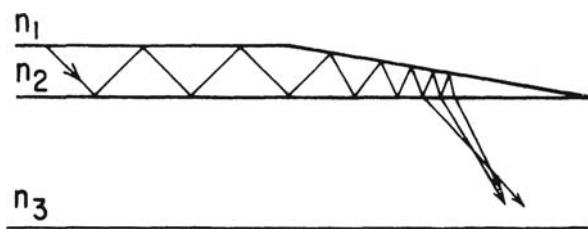


Fig. 7.11 Diagram of a tapered coupler

7.6 Tapered Mode Size Converters

The tapered couplers described in Section 7.6 relied on a reduction in waveguide cross-sectional dimensions sufficient to make the waveguide go below cut off, so that energy radiated out of the guide. However, if the dimensions of the waveguide are never reduced below the cutoff level for the particular mode in use, no energy is radiated and tapered waveguide structures can be used as mode size converters. Many different geometries can be used to either expand or contract the mode. A review of various taper designs that have been used has been done by Moerman et al. [26]. Some of the lateral tapers that they considered are shown in Fig. 7.12.

The functioning of the mode size converters shown in Fig. 7.12 is evident in the various diagrams. Mode size is either increased or decrease according to the shape of the taper and the direction of propagation. One general rule that should be followed in the design of such tapers is that the taper angle in the transition region is sufficiently small to prevent coupling of power from the fundamental mode into the higher order taper modes. It is also important that the minimum cross-sectional dimensions of the taper not be so small that the waveguide goes below cutoff, unless it is desired that optical energy be transferred to another waveguide as in the case of Fig. 7.12c. There, in the tapered section, the width of the upper ridge is sufficiently decreased so that the optical mode in the upper ridge is in cutoff near the taper facet and only a wide optical mode, defined by the broad lower ridge, is supported. The tapers shown in Fig. 7.12 are all lateral tapers, which can be defined by standard photolithographic techniques. Vertical tapers, in which the thickness of the waveguiding layer is changed, can also be used to make mode size converters. However,

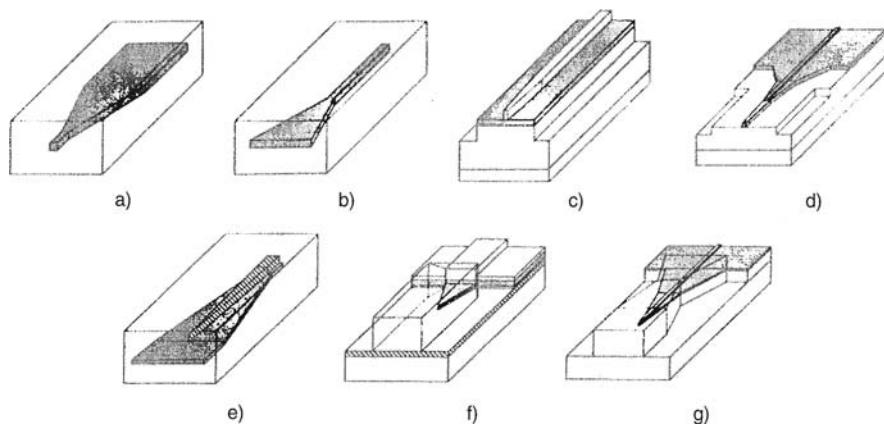


Fig. 7.12 a–g Lateral taper designs. **a** Lateral down-tapered buried waveguide. **b** Lateral up-tapered buried waveguide. **c** Single lateral taper transition from a ridge waveguide to a fiber-matched waveguide. **d** Multisection taper transition from a ridge waveguide to a fiber-matched waveguide. **e** Dual lateral overlapping buried waveguide taper. **f** Dual lateral overlapping ridge waveguide taper. **g** Nested waveguide taper transition from a ridge waveguide to a fiber-matched waveguide [26] ©1997 IEEE

special growth and etching techniques are required to change the thickness along the taper [26].

7.7 Fiber to Waveguide Couplers

A key element of an integrated optic system is the coupler that transfers the optical wave between a fiber, used for long distance transmission, and a waveguide of the OIC used for signal processing. Research in recent years has resulted in the development of a number of different types of fiber-to-waveguide couplers.

7.7.1 Butt Coupling

The fiber may be directly butted in contact with the waveguide, without any interfacing device, in an end-on alignment. If the cross-sectional area of the fiber core and the waveguide are closely matched, high efficiency coupling can be achieved, as in the case of butt-coupled channel waveguides or a laser diode and a channel waveguide [26]. An index matching fluid can be used to reduce reflection loss at the interface. The greatest problem with the butt coupling approach is that it is extremely difficult to establish and maintain correct alignment, since both the fiber core and waveguide typically have micron sized dimensions.

A mechanical arrangement for effectively coupling a single-mode fiber to a laser diode (or waveguide) which has been demonstrated by Enoch [4] is shown in Fig. 7.13. The coupling problem is difficult because the fiber core diameter is only a few micrometers and the laser light emitting region is less than 1 micrometer thick. Also, it is necessary to maintain the alignment while the fiber is permanently bonded in place. Piezoelectrically driven micromanipulators can be used to align the fiber, as mentioned previously in this chapter. To provide permanent bonding of the fiber and laser, the fiber is metallized and soldered to a fiber mount substrate. That substrate is, in turn, attached to a thermoelectric heat sink with solder. On the underside of the fiber mount substrate (shown in the upper left corner of Fig. 7.13) are metallization leads and a thin film resistor which are used to melt the solder by passage of an electrical current. A computerized feedback control system is used to move the laser and fiber into optimum alignment while current passing through the thin film resistor keeps the solder fluid. When the desired alignment has been achieved, the computer turns off the current, allowing the solder to cool and set. Thus the laser and fiber remain permanently bonded to the thermoelectric heat sink and alignment is maintained.

More recently, the silicon V-groove and flip-chip techniques, which have been used for years in the electrical integrated circuits industry [28], have been applied to the problem of aligning a channel waveguide with a single-mode fiber. Sheem et al. [29] have used a Si V-groove/flip-chip coupler to couple a 3 μm wide diffused LiNbO₃ waveguide to a 4.5 μm core single-mode fiber. They obtained a coupling

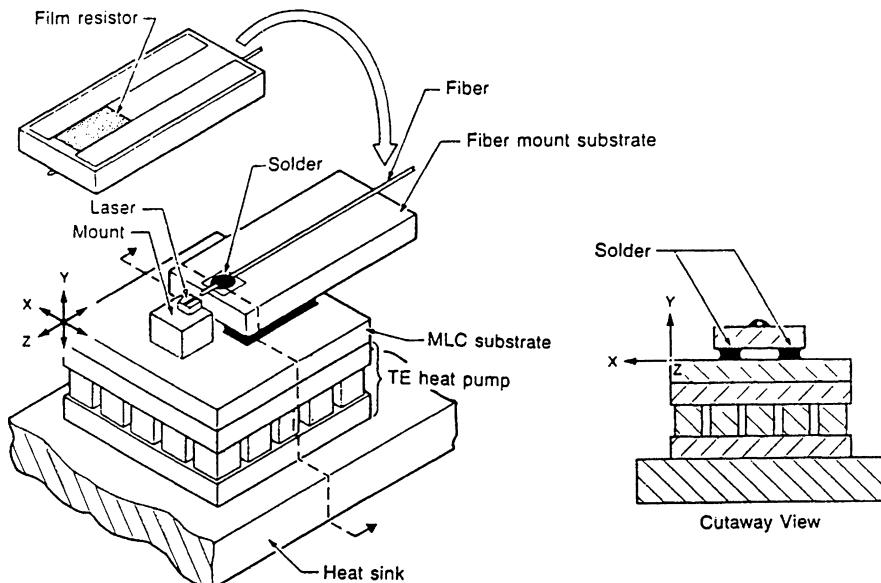


Fig. 7.13 Optical fiber interconnect to a single-mode laser diode [4] (Diagram courtesy of Tektronix, Inc.)

efficiency of 75% into both TE and TM modes. Murphy and Rice [30] have used the overlap between a precision-etched V-groove silicon chip and the top surface of a Ti-diffused LiNbO_3 waveguide substrate to simultaneously align an array of single mode fibers with corresponding channel waveguides. The coupling structure, shown in Fig. 7.14, aligns all but one of the six degrees of freedom automatically. Only the transverse alignment in the x direction must be adjusted for maximum coupling. Although only five fibers are shown in the diagram for simplicity, the method has been used to attach arrays of twelve fibers with average excess loss of 0.9 dB at a wavelength of $\lambda_0 = 1.3 \mu\text{m}$. More recent work by Sugita et al. [31] has reduced loss to 0.6 dB in an 8 groove interface.

Etched grooves can also be utilized to align optical fibers with either planar or channel waveguides in polymers [32]. Precisely oriented and sized slots are generated by eximer-laser ablation to accept the core of an optical fiber from which the cladding has been stripped. Angular precision of 0.3° and translational accuracy of $0.5 \mu\text{m}$ have been achieved, resulting in coupling losses less than 0.5 dB.

Chung et al. [33] have developed a technique for coupling multichannel waveguides to a fiber array using V-grooves in Perspex® acrylic substrates that are accurately machined by a computer-controlled CO_2 laser system. An excess loss of less than 1 dB was reported for coupling between 1×2 , 1×3 , and 1×4 branching circuits and an array of N-port fibers.

The optimization of coupling efficiency has remained a topic of continuing interest. Detailed theoretical models have been developed to describe the coupling

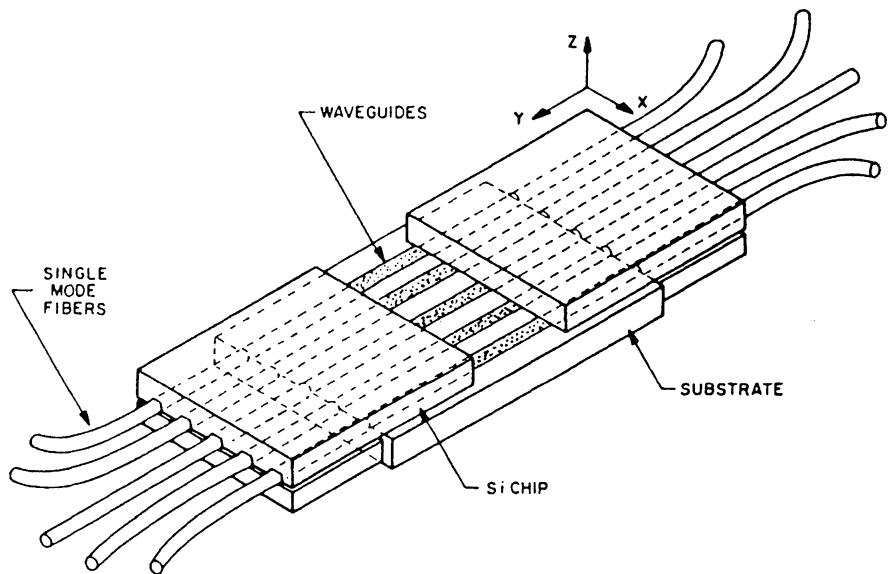


Fig. 7.14 Structure for self-aligned coupling of fibers and waveguides [30]

between optical fibers and both planar [34] and channel [35] waveguides, and mode-shape adapters have been implemented [36, 37] to improve the match between circular cross-section fibers and rectangular waveguides.

New micromachining techniques have been developed to improve alignment, as shown in Fig. 7.15 [38].

Standard processing techniques shown in (a) call for segregation of the fabricated rib waveguide device and the alignment structure while also having to saw cut and

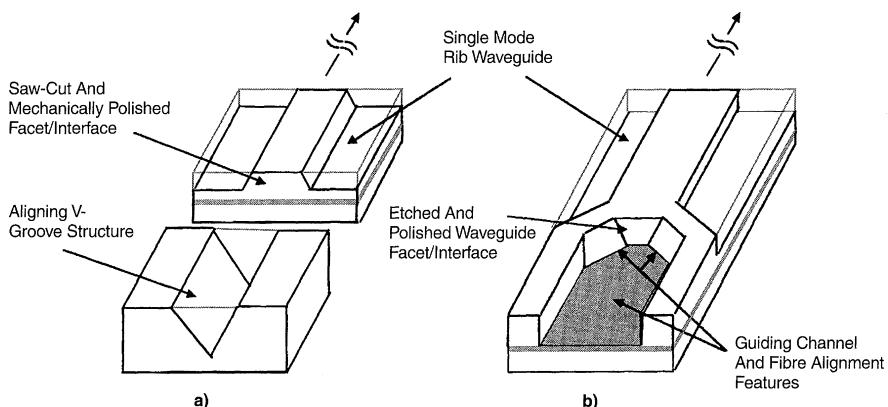


Fig. 7.15 a,b Comparison of a standard processing technique with b micromachined fiber alignment process [38]

mechanically polish each waveguide facet in order to achieve an optically smooth surface. In the micromachining process shown in (b) fabrication of both fiber alignment features and optical quality waveguide facet polishing are completed concurrently and integrated on the same substrate. The combinational advantages of this processing technique greatly reduce the power loss through misalignment errors and surface roughness on the waveguide sidewall or end facets.

To conclude this section on butt coupling of fibers, we note that light-emitting diodes (LED's), which emit over a relatively large surface area, can be coupled to optical fibers with comparative ease. For a thorough discussion of this topic, the reader is referred to a review by Barnoski [39].

7.7.2 High Density Multifiber Connectors

As the use of optical fiber interconnects has become wide spread in telecom and datacom systems the demand for multifiber connectors that are suitable for field applications has become very strong. To couple large numbers of single mode fibers repeatedly with minimal losses is a daunting task. However, clever solutions to the problem have been devised. For example, Takaya et al. [40] have designed mechanically transferable (MT) and multifiber push-on (MPO) type connectors that connect as many as 60 fibers. The basic configuration of two 60-fiber ribbon connectors is shown in Fig. 7.16.

The connectors were constructed by using a monolithic 60-fiber ferrule which aligns five stacked 12-fiber ribbons. The fiber cores are arranged in 5 rows of 12 at intervals of 250 μm . These connectors were designed to have the same outer dimensions as conventional MT type and MPO type connectors. The connectors have a high packaging density of 286 fibers/cm² for the MT type and 40 fibers/cm² for the MPO type. A key feature of these connectors that permits accurate alignment of all of the fibers is the pair of guide holes that closely fit a matching pair of guide pins on the mating connector as shown in Fig. 7.17.

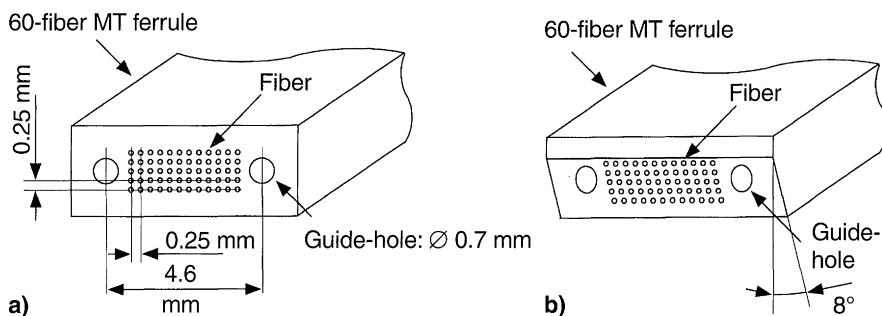


Fig. 7.16 a,b Endface geometries of 60-fiber connectors **a** MT type **b** MPO type [40] © 1999 IEEE

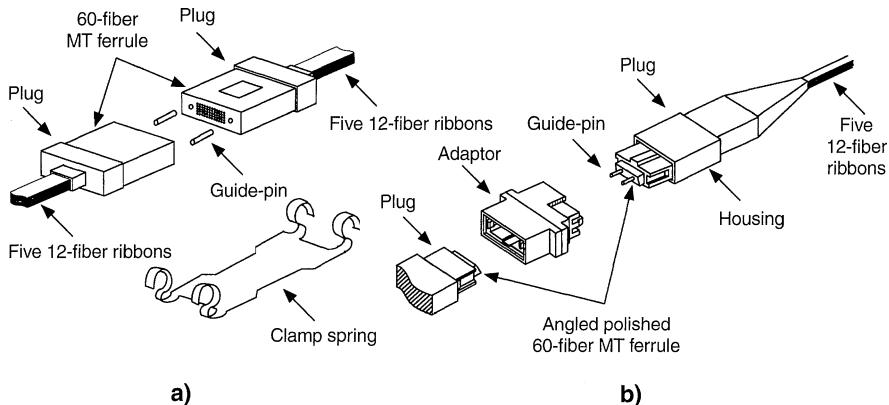


Fig. 7.17 a,b Structure of 60 fiber connectors **a** MT type **b** MPO type [40] © 1999 IEEE

The insertion losses of the 60-fiber MT type connector were measured using an LED operating $1.3 \mu\text{m}$. The average insertion loss was 0.20 dB with a maximum value of 1.00 dB . The average insertion loss of the 60-fiber MPO type connector was 0.70 dB with a maximum value of 1.80 dB . The maximum loss increase was less than 0.5 dB for 200 reconnections.

Problems

- 7.1 A strontium titanate prism ($n_p = 2.32$) is used as an output coupler to couple light out of a Ta_2O_5 waveguide ($n_2 = 2.09$). Three “ m -lines” are visible at angles of 36.5° , 30.2° , and 24.6° from the waveguide surface. The output face of the prism makes an angle of 60° with the waveguide surface and the wavelength λ_0 is 9050 \AA . What are the β 's of the three modes?
- 7.2 If a rutile prism ($n_p = 2.50$) is used as an input coupler to the same waveguide as in Problem 7.1, what angle should the incident light beam make with the waveguide surface to efficiently couple into the lowest order mode? Assume the input face of the prism makes an angle of 60° with the waveguide surface.
- 7.3 In the case of a prism coupler, explain why the condition

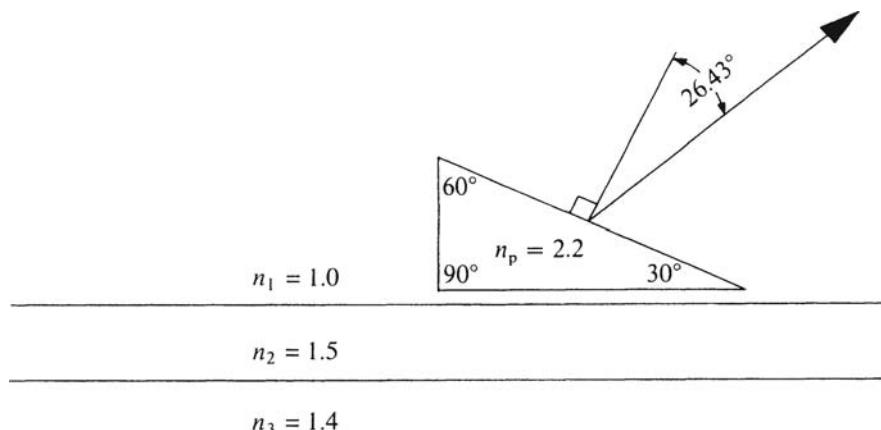
$$\beta_m = \frac{2\pi n_p}{\lambda_0} \sin \theta_m,$$

must be satisfied in order for coupling to occur. Derive the equation from geometric considerations (n_p : index of refraction of prism, β_m : phase constant of the mode in the waveguide, θ_m : angle between normal to waveguide surface and incident ray of light within the prism, and λ_0 : wavelength of light in air).

- 7.4 A strontium titanate right-angle prism ($n_p = 2.32$) is used as an input coupler to couple light of 9050 \AA (vacuum) wavelength into a Ta_2O_5 waveguide

($n = 2.09$) from an air beam. The face of the prism makes an angle of 45° with the surface of the waveguide. The β of the lowest-order mode in the waveguide is $1.40 \times 10^5 \text{ cm}^{-1}$.

- (a) Make a sketch of the coupler labeling the significant angles ϕ_1 , ϕ_2 , ϕ_3 , etc. Use these labels in your calculations in part b.
 - (b) What angle must the input laser beam (in air) make with the surface of the waveguide in order to be coupled into the lowest-order mode?
- 7.5 A grating situated on a planar waveguide can act as a 180° reflector for the waves within the guide. If the propagation constant of the guided mode is $\beta = 1.582 k$ and $\lambda_0 = 0.6328 \mu\text{m}$, find the smallest grating spacing A that will cause the mode to be reflected.
- 7.6 A rutile prism ($n_p = 2.50$) is used to couple light of vacuum wavelength $\lambda_0 = 0.9050 \mu\text{m}$ to the fundamental mode in a waveguide which has a refractive index $n_g = 2.09$. Given that the phase constant of the fundamental mode is $\beta_0 = 1.44 \times 10^5 \text{ cm}^{-1}$ what angle γ must the input face of the prism make with the waveguide surface and what angle ϕ must the optical beam within the prism make with the waveguide surface in order to obtain the most efficient coupling.
- 7.7 A grating with spacing $\Lambda = 0.4 \mu\text{m}$, situated on a GaAs planar waveguide, is to be used for coupling a beam of light from a He-Ne laser ($\lambda_0 = 1.15 \mu\text{m}$) into the waveguide. If the propagation constant for the lowest-order mode in the guide is $\beta_0 = 3.6 k$, what angle must the laser beam make with the surface of the waveguide in order to couple to this mode? Assume first-order coupling, i.e. $|v| = 1$.
- 7.8 A thin film waveguide has $n_1 = 1$, $n_2 = 1.5$ and $n_3 = 1.462$. The waveguide thickness is $0.9 \mu\text{m}$. Light from a He-Ne laser ($\lambda_0 = 6328 \text{ \AA}$) is being guided in the (fundamental) TE_0 mode, for which the effective refractive index of the guide is $n_{\text{eff}} = 1.481$. If a $45^\circ - 45^\circ - 90^\circ$ prism with index $n_p = 2.25$ is used as an output coupler, what angle will the exiting beam make with the surface of the waveguide?
- 7.9 A prism coupler with index $n_p = 2.2$ is used to observe the modes of a waveguide as shown below. The light source is a He-Ne laser with $\lambda_0 = 6328 \text{ \AA}$. If the light from a particular mode is seen at an angle of 26.43° with the normal to the prism surface, what is the propagation constant β for that mode?
- 7.10 A thin-film waveguide is characterized by the following parameters: $n_1 = 1$, $n_2 = 1.5$ and $n_3 = 1.462$. The waveguide thickness = $0.9 \mu\text{m}$. The fundamental TE mode of the waveguide is excited by coupling the beam from a He-Ne laser ($\lambda_0 = 6328 \text{ \AA}$) through a prism having $n_p = 2.25$.
- (a) What is the required angle of incidence of the light at the interface between the prism and the waveguide (inside the prism) if the effective index of refraction for the above mode is $n_{\text{eff}} = 1.481$?
 - (b) For the mode and waveguide described above, what is the bounce angle if the mode is represented by the ray optic method?



- 7.11 Can a grating coupler that is used to couple light of $0.85 \mu\text{m}$ into a given waveguide also be used to couple light of $1.06 \mu\text{m}$ into the same waveguide? Explain how.
- 7.12 If a quartz prism with an index of refraction = 1.65 is used to fabricate a grating on a GaAs waveguide as shown in Fig. 7.10, with the angle $\alpha = 5^\circ$:
- What is the spacing of the grating bars if a He-Cd laser (wavelength = 325 nm) is used?
 - What is the spacing of the grating bars if an ArF laser (wavelength = 193 nm) is used?

References

- A. Yariv: IEEE J. **QE-9**, 919 (1973)
- R.G. Hunsperger, A. Yariv, A. Lee: Appl. Opt. **16**, 1026 (1977)
- J.M. Hammer, D. Botez, C.C. Ncil, J.C. Connolly: J. Appl. Phys. **39**, 943 (1981)
- S. Enoch: Optical fiber interconnect to a single mode laser. OSA/IEEE Meeting on Integrated and Guided Wave Optics, Atlanta, GA (1986)
- M. Yanagisawa, H. Terui, K. Shuto, T. Miya, M. Kobayashi: Photon. Techn. Lett. **4**, 21 (1992)
- M. Aoki, M. Suzuki, T. Taniwataru, H. Sano, T. Kawano: Microwave Opt. Techn. Lett. **7**, 132 (1994)
- T. Tamir: Beam and waveguide couplers, in *Integrated Optics*. T. Tamir (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979) Chap. 3, in particular, pp. 102–107
- M.S. Klimov, V.A. Sychugov, A. Tishchenko, O. Parriaux: Fiber Integr. Opt. **11**, 85 (1992)
- L. de S. Menezes, A. Mazzei, S. Götzinger, O. Benson, V. Sandoghdar: Optimizing the coupling of light via a prism to high-Q modes of a microsphere resonator using a near-field probe, Proc. Encontro Nacional de Fisica de Materia Condensada, ENFMC XXIX, Sao Lourenco, Brazil, May 2006
- P.K. Tien: Appl. Opt. **10**, 2395 (1971)
- T. Tamir, S.T. Peng: Appl. Phys. **14**, 235 (1977)

12. M. Shams, D. Botez, S. Wang: Opt. Lett. **4**, 96 (1979)
13. A. Gruss, K.T. Tam, T. Tamir: Appl. Phys. Lett. **36**, 523 (1980)
14. K. Rokushima, J. Yamakita: J. Opt. Soc. Am. **73**, 901 (1983)
15. K.C. Chang, T. Tamir: Appl. Opt. **19**, 282 (1980)
16. K.C. Chang, V. Shah, T. Tamir: J. Opt. Soc. Am. **70**, 804 (1980)
17. S. Somekh, E. Garmire, A. Yariv, H. Garvin, R.G. Hunsperger: Appl. Opt. **12**, 455 (1973)
18. M.H. Lim, T.E. Murphy, J. Ferrera, J.N. Damask, H.I. Smith: Fabrication techniques for grating-based optical devices, J. Vac. Sci. Technol. B: Microelectron. Nanometer Struc. **17**, 3208 (1999)
19. M. Dakss, L. Kuhn, P.F. Heidrich, B.A. Scott: Appl. Phys. Lett. **16**, 523 (1970)
20. E. Kapon, A. Katzir: J. Appl. Phys. **53**, 1387 (1982)
21. Yu.I. Ostrovsky, M.M. Butusov, G.V. Ostrovskaya: *Interferometry by Holography*, Springer Ser. Opt. Sci., Vol. 20 (Springer, Berlin, Heidelberg 1980)
22. Yu. I. Ostrovsky, V.P. Shchepinov, V.V. Yakovlev: *Holographic Interferometry in Experimental Mechanics*, Springer Ser. Opt. Sci., Vol. 60 (Springer, Berlin, Heidelberg 1991)
23. H. Yen, M. Nakamura, E. Garmire, S. Somekh, A. Yariv: Opt. Commun. **9**, 35 (1973)
24. S. Hava, H.B. Sequeira, R.G. Hunsperger: Fabrication of monolithic Peltier cooling structures for semiconductor laser Diodes. Joint Meeting, Nat'l Sci. Foundation, Grantee-User Group in Opt. Commun. and the Opt. Nat'l Telecom-mun. and Inform. Administr. Task Force on Opt. Commun., St. Louis, MO (1981) Proc. pp. 46–51
25. P.K. Tien, R.J. Martin: Appl. Phys. Lett. **18**, 398 (1974)
26. I. Moerman, P.P. Van Daele, P.M. Demeester: A review on fabrication technologies for the monolithic integration of tapers with III-V semiconductor devices. IEEE J. Selected topics in Quantum Elect. **3**, 1308 (1997)
27. W.L. Emkey: IEEE J. LT-**1**, 436 (1983)
28. A.B. Glaser, G.E. Subak-Sharpe: *Integrated Circuit Engineering* (Addison-Wesley, Reading, MA 1977) pp. 263–265 and 267–268
29. S.K. Sheem, C.H. Bulmer, R.P. Moeller, W.K. Burns: High efficiency single-mode fiber/channel waveguide flip-chip coupling. OSA Topical Meeting on Integrated Optics, Incline Village, NV (1980)
30. E.J. Murphy, T.C. Rice: IEEE J. **QE-22**, 928 (1986)
31. A. Sugita, K. Onosa, Y. Ohnori, M. Yasu: Fiber Integr. Opt. **12**, 347 (1993)
32. B.L. Booth: Optical interconnection polymers, in *Polymers for Lightwave and Integrated Optics: Technology and Applications*, L.A. Hornak (ed.) (Dekker, New York 1992)
33. P.S. Chung, W.Y. Hung, H.P. Chan: Fabrication of waveguide - fiber array couplers using laser-machined V-groove techniques in perspex substrates, Microwave Opt. Technol. Lett. **2**, 421 (1989)
34. A.T. Andreev, K.P. Panajotov, B.S. Zatirova, J.B. Koprinarova: SPIE Proc. **1973**, 72 (1993)
35. J. Lee, H. Lee, C. Lee: SPIE Proc. **1813**, 76 (1991)
36. T. Brenner, H. Melchior: IEEE Photon. Techn. Lett. **5**, 1059 (1993)
37. M. Mashayekhi, W.J. Wang, S.I. Najafi: Semiconductor device to optical fiber coupling using low-loss glass taper waveguide. Opt. Eng. **36**, 3476 (1997)
38. M.A. Rosa, N.Q. Ngo, D. Sweatman, S. Dimitrijev, H.B. Harrison: Self-alignment of optical fibers with optical quality end-polished silicon rib waveguides using wet chemical micromachining techniques. IEEE J. Selected topics in Quantum Elect. **5**, 1249 (1999)
39. M.K. Barnoski: Fiber couplers, in *Semiconductor Devices for Optical Communications*, H. Kressel (ed.) 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982) pp. 201–211
40. M. Takaya, S. Nagasawa, Y. Murakami: Design and performance of very high-density multi-fiber connectors employing monolithic 60-fiber ferrules, IEEE Phot. Technol. Lett. **11**, 1446 (1999)

Chapter 8

Coupling Between Waveguides

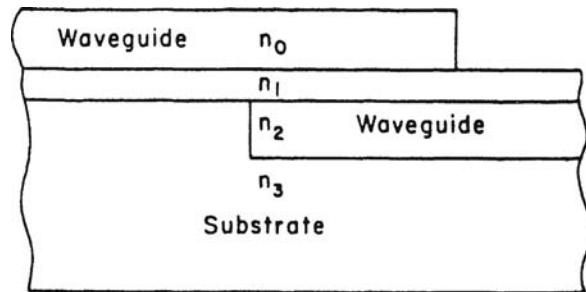
The phenomenon of optical tunneling can be used not only to couple energy from a fiber or a beam to a waveguide, as described in Chapter 7, but also to couple one waveguide to another. Couplers of this type are usually called directional couplers because the energy is transferred in a coherent fashion so that the direction of propagation is maintained. Directional couplers have been fabricated in two basic geometries: multilayer planar structures, and dual side-by-side channel waveguides. In this chapter the different types of waveguide to waveguide couplers are described and a concise theory of operation is developed. For a thorough mathematical treatment of these devices the reader is referred to the work of Burns and Milton [1].

8.1 Multilayer Planar Waveguide Couplers

While butt coupling can be used to couple two planar waveguides, as mentioned in Chapter 7, the more common method is to bring the guides into close proximity and allow coupling to occur through phase coherent energy transfer (optical tunneling), as shown in Fig. 8.1. The indices n_0 and n_2 in the guiding layers must be larger than n_1 and n_3 , and the thickness of the confining layer 1 must be small enough that the evanescent tails of the guided modes overlap. In order for energy transfer to occur between the two guides, they must have identical propagation constants. Thus, the indices and the thicknesses of the waveguiding layers must be very carefully controlled to provide matching propagation constants. This is difficult to do but has been accomplished with excellent results. As in the case of other devices embodying the synchronous coupling principle, such as the prism coupler, the interaction length must be carefully chosen for optimum coupling. The condition for total transfer of energy is again given by (7.10), but the value of κ would obviously be different for a pair of overlapped planar waveguides than for a prism coupler.

This method of coupling is difficult to use with deposited thin film waveguides because thicknesses and indices of refraction cannot be conveniently controlled. However, in the case of epitaxially grown waveguides, the superior control of thicknesses and index (by controlling composition) make the problem of matching propagation constants much easier to solve. The method seems to be particularly well

Fig. 8.1 Coupling between two planar waveguides by optical tunneling. Transfer of energy occurs by phase coherent synchronous coupling through the isolation layer with index n_1

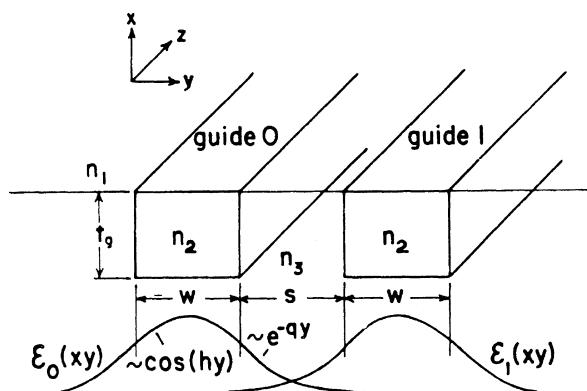


suited to coupling an integrated laser diode to a waveguide. For example Vawter et al. [2] have used a multilayer planar structure to couple a GaAlAs transverse junction stripe (TJS) laser to a waveguide of the same material. In that case the intervening isolation layer was composed of lower refractive index GaAlAs. The method has also been used by Utaka et al. [3] to couple lasers and waveguides composed of GaInAsP, separated by a thin layer of InP. This twin-guide structure can be used to fabricate an efficient single-mode laser with distributed feedback gratings [4]. A very efficient laser can be obtained by using two separate, but coupled, waveguides to isolate the active region, where photons are generated, from the distributed feedback region. Such devices will be discussed in more detail in Chapter 13.

8.2 Dual-Channel Directional Couplers

The dual-channel directional coupler, which is analogous to the microwave dual-guide multihole coupler [5], consists basically of parallel channel optical waveguides sufficiently closely spaced so that energy is transferred from one to the other

Fig. 8.2 Diagram of dual-channel directional coupler. The amplitudes of the electric field distributions in the guides are shown below them



by optical tunneling, as shown in Fig. 8.2. This energy is transferred by a process of synchronous coherent coupling between the overlapping evanescent tails of the modes guided in each waveguide. Photons of the driving mode, say in guide 0, transfer into the driven mode in guide 1, maintaining phase coherence as they do. This process occurs cumulatively over a significant length; hence, the light must propagate with the same phase velocity in each channel in order for this synchronous coupling to occur. The fraction of the power coupled per unit length is determined by the overlap of the modes in the separate channels. Thus, it depends on the separation distance s , the interaction length L , and the mode penetration into the space between channels, which can be characterized by the extinction coefficients p and q . (See Chapters 2 and 3 for a discussion of the mode shapes.)

8.2.1 Operating Characteristics of the Dual-Channel Coupler

In a dual-channel coupler, the energy transfers alternately from one waveguide to the other, and then back again if the interaction length is sufficient. If one were to measure the optical energy density while moving in the z direction along one channel of a direction coupler, a sinusoidal variation with distance would be observed. For a coupler to transfer any given fraction of the energy, it is necessary only to bend away the secondary channel at the proper point. In this way, for example, either a 10 dB coupler for measurement padding, a 3 dB coupler for beam splitting, or a 100% coupler for beam switching can be made.

The experimentally measured transmission characteristics of a dual-channel 3 dB coupler are shown in Fig. 8.3. The diagram at the left shows the position of the input laser beam, and the photograph on the right shows an oscilloscope scan

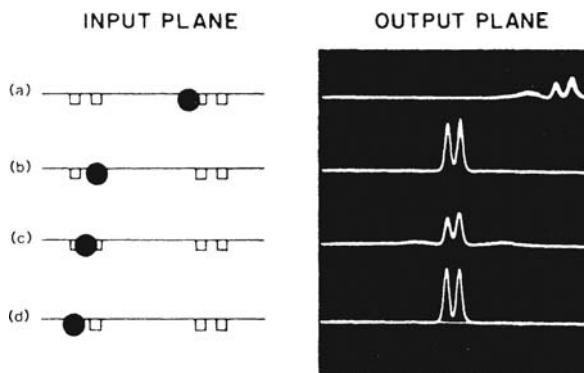


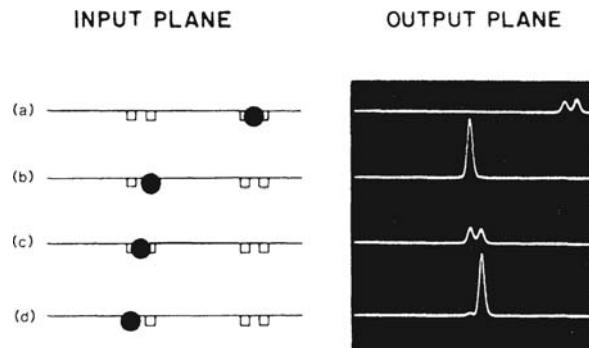
Fig. 8.3 Optical power distribution at the output of a 3 dB dual-channel directional coupler, for various input conditions (as explained in the text). The oscilloscopes of output power were made using a scanning system like that shown in Fig. 2.3. The waveguides, which were formed by proton bombardment of the GaAs substrate, had $3 \mu\text{m} \times 3 \mu\text{m}$ cross-section and were separated by $3 \mu\text{m}$. The interaction length was 1 mm.

(in the y direction) of the optical power densities in the waveguides at the output plane. The waveguides forming this particular coupler were fabricated by 300 keV proton bombardment of a GaAs substrate [6]. The waveguides had a 3 μm square cross-section and were separated by approximately 3 μm . A 1.15 μm wavelength He-Ne laser beam was focused onto the cleaved input plane of the sample, which happened to contain two pairs of coupled channel waveguides. When the laser beam was focused directly on one guide, as in cases b and d of Fig. 8.3, the optical energy at the output plane was equally divided between both of the coupled guides, i.e. a 3 dB coupling was evident. When the laser beam covered only a portion of one of the guides, as in case a, or a portion of both guides as in case c, a reduced amount of light was coupled, and appeared at the output still roughly equally divided between the two guides of the dual-channel pair.

The coupling effect is even more striking in the transmission characteristics of a 100% coupler, as shown in Fig. 8.4. This device was identical to the 3 dB coupler except that it was 2.1 mm long rather than 1 mm. This length was chosen to produce total transfer of energy at the output plane. In this case, when the input laser beam was focused on either channel of a dual pair, as in (b) and (d), the light was contained in the opposite channel of the pair at the output plane. When the input beam was spread across both channels, as in (a) or (c), a reduced amount of light was coupled into both channels.

The results shown in Figs. 8.3 and 8.4 demonstrate that dual-channel couplers can produce a very efficient transfer of optical energy from one guide to another over a relatively short interaction length. The embedded waveguides produced by ion implantation, which lie beneath the surface of the substrate and have an index difference ~ 0.005 from the substrate material, are particularly useful for dual-channel coupler fabrication, because the mode tails extend significantly into the space between guides. This is not the case, for example, with a ridged guide formed on the surface of the substrate and surrounded on three sides by air. Nevertheless, effective dual-channel couplers can be made in glass if the critical dimensions and indices are carefully controlled [7]. Various techniques of fabricating dual-channel

Fig. 8.4 Optical power distribution at the output of a 100% dual-channel directional coupler for various input conditions. The waveguides were like those of Fig. 8.3, except the interaction length was 2.1 mm



couplers will be described and compared in more detail in a later section of this chapter, following a discussion of the theory of operation.

8.2.2 Coupled-Mode Theory of Synchronous Coupling

A concise theory of operation of the dual-channel directional coupler can be developed by following the coupled mode theory approach of Yariv [8, 9]. This model has been compared with experimental measurements of three-arm couplers fabricated in Ti:LiNbO₃ by Peall and Syms [10], and excellent qualitative agreement was obtained. The electric field of the propagating mode in the waveguide is described by

$$\bar{E}(x, y, z) = A(z)\bar{\mathcal{E}}(x, y), \quad (8.1)$$

where $A(z)$ is a complex amplitude which includes the phase term $\exp(-i\beta z)$. The term $\bar{\mathcal{E}}(x, y)$ is the solution for the field distribution of the mode in one waveguide, assuming that the other waveguide is absent. By convention, the mode profile $\bar{\mathcal{E}}(x, y)$ is assumed to be normalized to carry one unit of power. Thus, for example, the power in guide number 1 is given by

$$P_1(z) = |A_1(z)|^2 = A_1(z)A_1^*(z). \quad (8.2)$$

The coupling between modes is given by the general coupled mode equations for the amplitudes of the two modes. Thus,

$$\frac{dA_0(z)}{dz} = -i\beta_0 A_0(z) + \kappa_{01} A_1(z), \quad (8.3)$$

and

$$\frac{dA_1(z)}{dz} = -i\beta_1 A_1(z) + \kappa_{10} A_0(z), \quad (8.4)$$

where β_0 and β_1 are the propagation constants of the modes in the two guides, and κ_{01} and κ_{10} are the coupling coefficients between modes.

Consider the guides shown in Fig. 8.2. Assume that the guides are identical and that they both have an exponential optical loss coefficient α . Thus,

$$\beta = \beta_r - i\frac{\alpha}{2}, \quad (8.5)$$

where $\beta = \beta_0 = \beta_1$, and β_r is the real part of β . For the case of identical guides, it is obvious from reciprocity that

$$\kappa_{01} = \kappa_{10} = -i\kappa, \quad (8.6)$$

where κ is real. Then, using (8.5) and (8.6), Eqs. (8.3) and (8.4) can be rewritten in the form

$$\frac{dA_0(z)}{dz} = -i\beta A_0(z) - i\kappa A_1(z), \quad (8.7)$$

and

$$\frac{dA_1(z)}{dz} = -i\beta A_1(z) - i\kappa A_0(z). \quad (8.8)$$

If it is assumed that light is coupled into guide 0 at the point $z = 0$, so that boundary conditions for the problem are given by

$$A_0(0) = 1 \quad \text{and} \quad A_1(0) = 0, \quad (8.9)$$

then the solutions are described by

$$A_0(z) = \cos(\kappa z)e^{-i\beta z} \quad (8.10)$$

and

$$A_1(z) = -i \sin(\kappa z)e^{-i\beta z}. \quad (8.11)$$

Thus, the power flow in the guides is given by

$$P_0(z) = A_0(z)A_0^*(z) = \cos^2(\kappa z)e^{-\alpha z} \quad (8.12)$$

and

$$P_1(z) = A_1(z)A_1^*(z) = \sin^2(\kappa z)e^{-\alpha z}. \quad (8.13)$$

From (8.12) and (8.13), it can be seen that the power does indeed transfer back and forth between the two guides as a function of length, as shown experimentally in Figs. 8.3. and 8.4. Note also, in (8.10) and (8.11), the distinct phase difference that exists between the amplitude of the fields in the two guides. The phase in the driven guide always lags 90° behind the phase of the driving guide. Thus, initially at $z = 0$, the phase in guide 1 lags 90° behind that in guide 0. That lagging phase relationship continues for increasing z , so that at a distance z that satisfies $\kappa z = \frac{\pi}{2}$, all of the power has been transferred to guide 1. Then, for $\frac{\pi}{2} \leq \kappa z \leq \pi$, the phase in guide 0 lags behind that in guide 1, and so on. This phase relationship results from the basic mechanism which produces the coherent transfer of energy. The field in the driving guide causes a polarization in the dielectric material which is in phase with it, and which extends into the region between guides because of the mode tail. This polarization then acts to generate energy in the mode of the driven guide. It is a basic principle of field theory that generation occurs when polarization leads

the field, while dissipation occurs when polarization lags the field [11]. Thus, the lagging field in the driven guide is to be expected. Because of this definite phase relationship, the dual-channel coupler is a directional coupler. No energy can be coupled into a backward wave traveling in the z direction in the driven waveguide. This is a very useful feature in many applications.

From (8.12) and (8.13), it can be seen that the length L necessary for complete transfer of power from one guide to the other is given by

$$L = \frac{\pi}{2\kappa} + \frac{m\pi}{\kappa}, \quad (8.14)$$

where $m = 0, 1, 2, \dots$. In a real guide, with absorption and scattering losses, β is complex. Hence, the total power contained in both guides decreases by a factor $\exp(-\alpha z)$. The theoretical power distribution as a function of z from (8.12) and (8.13) is plotted in Fig. 8.5.

The coupling coefficient κ is a strong function of the shape of the mode tails in the guides. For well confined modes, in which the overlapping of the tails causes only a negligible perturbation of the basic mode shape, it can be shown that the coupling coefficient is given by [12]

$$\kappa = \frac{2h^2qe^{-qs}}{\beta W(q^2 + h^2)}, \quad (8.15)$$

where W is the channel width, s is the separation, h and β are the propagation constants in the y and z directions, respectively, and q is the extinction coefficient in the y direction. It will be recalled that these parameters have been assumed to be identical for both waveguides.

In a practical situation, it may be difficult to fabricate two identical waveguides to form a coupler. If, for example, the guides do not have exactly the same thickness and width, the phase velocities will not be the same in both. This will not necessarily destroy the coupling effect entirely. If the difference in phase constants $\Delta\beta$ is small, it can be shown [6] that the power distributions in the two guides are given by

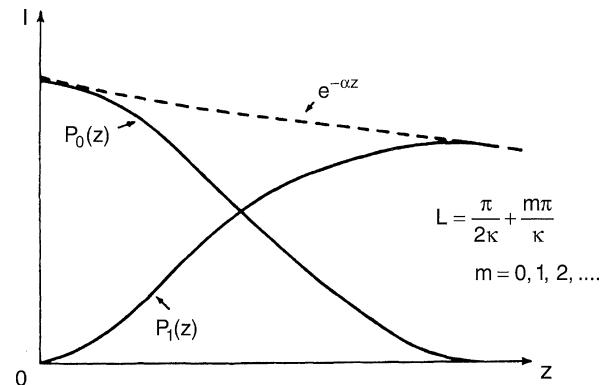


Fig. 8.5 Theoretically calculated power distribution curves for a dual-channel directional coupler. The initial condition of $P_0(0) = 1$ and $P_1(0) = 0$ has been assumed

$$P_0(z) \cos^2(gz)e^{-\alpha z} + \left(\frac{\Delta\beta}{2}\right)^2 \frac{\sin^2(gz)}{g^2} e^{-\alpha z}, \quad (8.16)$$

and

$$P_1(z) = \frac{\kappa^2}{g^2} \sin^2(gz)e^{-\alpha z}, \quad (8.17)$$

where

$$g^2 \equiv \kappa^2 + \left(\frac{\Delta\beta}{2}\right)^2. \quad (8.18)$$

It can be seen from (8.16), (8.17), (8.18) that, in the presence of a phase constant difference $\Delta\beta$, transfer of power will still occur. However, the transfer will be incomplete, since (8.16) has no zeros for any z .

The preceding equations can be used to calculate the expected performance of dual-channel directional couplers for the case of slightly non-identical guides, with a non-zero $\Delta\beta$. We will return to these equations in Chapter 9 to consider the case in which $\Delta\beta$ is made electrically controllable, to produce an optical modulator and switch.

8.2.3 Methods of Fabricating Dual-Channel Directional Couplers

One of the most convenient methods of fabricating the closely-spaced waveguides of the dual-channel directional coupler, in the case of GaAs (or GaP) substrates, is masked proton bombardment. The proton bombardment step is diagrammed in Fig. 8.6. A vapor-deposited gold mask, one or two micrometer thick, is used to block the protons in the regions surrounding the waveguides. In between the bars of the mask, the protons penetrate the semiconductor, producing a carrier-compensated waveguide as described in Chapter 4. A $1.5 \mu\text{m}$ thick layer of gold is sufficient to block 300 keV protons.

The mask is formed by first depositing a gold layer over the entire substrate surface; then applying a layer of photoresist on top of the gold by using the standard thin film spinning process described in Section 4.1.2. The photoresist is then exposed and developed, down to the gold using standard photolithographic techniques, to produce a striped pattern of resist with a ridged shape. This photoresist pattern is in turn used to mask the gold layer during ion-beam micromachining [13]. The ion-beam sputtering process can be easily controlled to remove the gold just down to the surface of the GaAs. If an ion such as Ar^+ or Kr^+ is used, the sputtering yield for gold will be relatively high, compared to that for either GaAs or photoresist. Thus, a $2 \mu\text{m}$ thick photoresist mask will adequately mask a $2 \mu\text{m}$ thick gold layer, and also very little GaAs will be removed, even if the process is not halted instantly when the GaAs surface is first reached. This is important because it

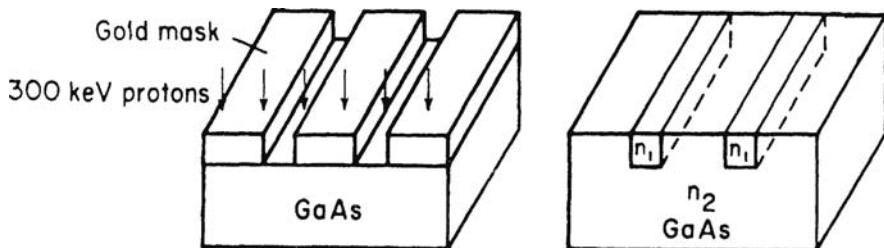


Fig. 8.6 Diagram of the masked proton implantation method of dual-channel coupler fabrication. For a proton energy of 300 keV, a $3 \mu\text{m}$ thick waveguide is formed in GaAs

provides a built-in compensation for any possible variations in thickness of the photoresist. Figure 8.7 shows a scanning-electron-microscope photograph of a mask formed on a GaAs substrate by the process which has been described. It can be seen that edge variations where the gold mask meets the GaAs are less than about $0.05 \mu\text{m}$, so that scattering loss should be negligible. Absorption loss in proton implanted waveguides is on the order of 1 cm^{-1} [14].

After the gold mask is in place on the substrate, channel waveguides are formed by proton implantation as described in Chapter 4. Couplers formed by 300 keV H^+ bombardment of GaAs to produce $2.5 \mu\text{m}$ wide guides spaced $3.9 \mu\text{m}$ apart, have exhibited a coupling coefficient κ equal to 0.6 mm^{-1} [15]. In this case, the increase in index of refraction within the guides due to proton bombardment Δn was calculated to be 0.0058. The coupling coefficient κ is a strong function of Δn . For example, a 10% reduction of Δn from the above value would increase κ to 0.74 mm^{-1} .

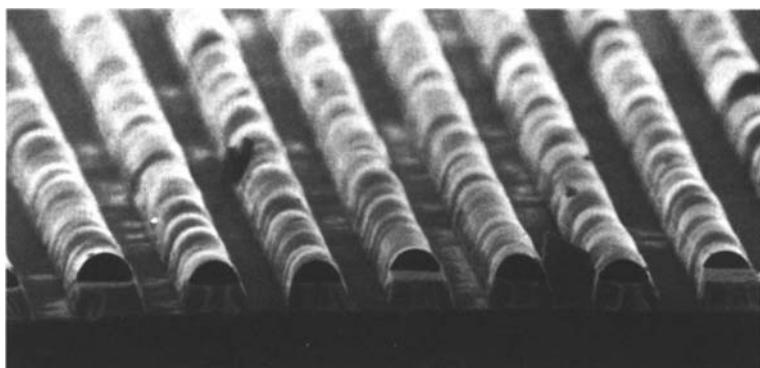


Fig. 8.7 Scanning-electron-microscope photograph of the evaporated gold mask used in coupler fabrication as in Fig. 8.6. Spacing between mask bars equals $3 \mu\text{m}$. (Remains of the photoresist mask can be seen above the gold stripes). (Photo courtesy of Hughes Research Laboratories, Malibu, CA)

As an alternative to proton bombardment, diffusion of a compensating dopant atom [16, 17] could be used to form the waveguides. Also, masked diffusion of a third element into a binary compound can be used to form a channel waveguide. For example, Martin and Hall [18] used an SiO_2 mask on ZnSe to form Cd diffused channel waveguides that were $10 \mu\text{m}$ wide. In LiNbO_3 and LiTiO_3 , diffusion of metal atoms can be used to produce the required waveguide pair. For example, Kondo et al. [19] have used Ti diffusion into LiNbO_3 to produce a waveguide modulator. In most cases, diffusion doping is much less predictable than ion implantation doping, exhibiting greater variations in dopant concentration profile. This makes diffusion doping more difficult to use for coupler fabrication, since the coupling coefficient is such a strong function of the index profiles in the guides. However, evanescent coupling between two indiffused waveguides has been mathematically modeled [17, 20].

Dual-channel couplers can also be fabricated by masking the substrate and depositing metallic strips on the surface to produce a closely spaced pair of strip-loaded guides. Campbell et al. [21] have used that approach to form a dual-channel coupler in GaAs. They began with an epitaxially grown, carrier-concentration-reduction type, planar waveguide on a heavily doped ($N_d \sim 10^{18} \text{ cm}^{-3}$) *n*-type substrate. Thin (5–25 μm wide) metallic strips were produced by using standard photoresist masking and etching techniques, and vapor depositing either Au, Ni or Pt. Separation between the waveguides was on the order of a few micrometers. This type of coupler is particularly useful in electro-optic modulator or switch applications, since the metal loading strips can also be used as electrodes to control the phase matching between the two channels. Devices of that type are discussed in more detail in Chapter 9.

All of the methods of dual-channel coupler fabrication that have been described thus far produce imbedded waveguides, which lie below the surface of the substrate. However, couplers can also be made by beginning with a substrate containing a planar waveguide over its entire surface and then removing superfluous regions by masking and etching to produce raised ridged channel guides. The planar waveguide is usually made by epitaxial growth. The masking is accomplished with standard photoresist techniques, and either chemical or ion-beam etching can be used. However, in general, chemical etchants are difficult to control to produce the required degree of sidewall smoothness. If ion-beam machining is used, the photoresist pattern is replicated in the waveguide layer just as it was in the gold layer used as a mask for proton bombardment, as previously described. Very smooth sidewalls can be produced using this method, as shown in Fig. 8.8. Edge variations are seen in the photograph to be less than a few hundred Angstroms. However, because of the larger index difference at the sidewall interface, scattering losses are particularly sensitive to wall smoothness. Hence, care must be exercised to minimize variations, limiting them to at most about 500 Å.

Another problem inherent in fabricating couplers with ridged channel guides is that the larger index difference confines the modes so well to their respective channels that insufficient coupling is obtained. To increase the coupling, only a partial removal of the material between channels can be performed, as shown in Fig. 8.9.

Fig. 8.8 Scanning electron microscope photograph of an ion-beam machined ridged channel waveguide in GaAs. The channel is 1.4 μm high and 2.0 μm wide [6]

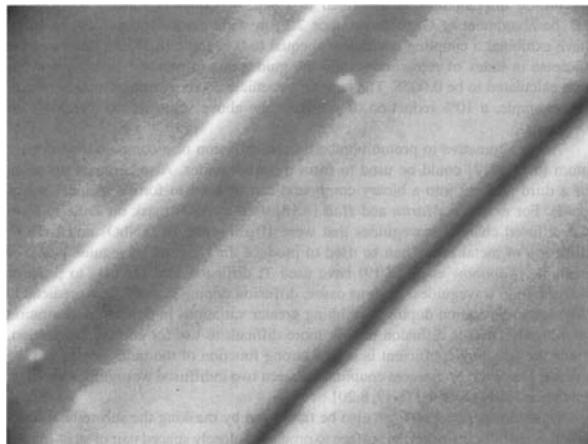
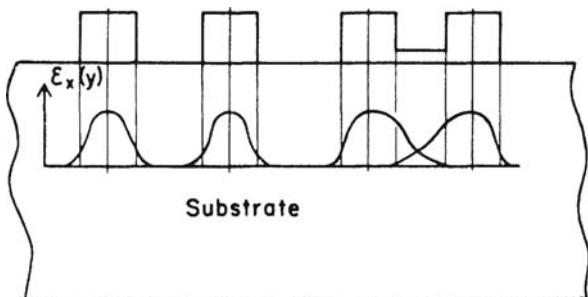


Fig. 8.9 Diagram of dual-channel couplers fabricated by masked etching of a planar waveguide layer. The electric field distributions are shown for the case of total etched isolation on the left and for bridged isolation on the right



The strength of coupling then depends on the thickness of the waveguide bridge between channels [6]. Alternatively, the coupling between two ridged channels can be increased by filling the region between them with another material having a slightly lower refractive index.

Ridged channel waveguides are most useful in materials such as LiNbO_3 and glasses, in which imbedded waveguides are difficult to produce. For example, Kaminow et al. [22] have used ion-beam machining to produce 19 μm wide ridged guides in LiNbO_3 . Kawabe et al. [23] have made ridged-waveguide directional couplers in LiNbO_3 by employing an ion-bombardment-enhanced etching technique. This method utilizes the fact that a material damaged by ion bombardment generally has a higher chemical etch rate than unbombarded material. The waveguide pattern was defined by masked 60 keV Ar^+ bombardment, and unwanted material was then removed from around the waveguides by etching in dilute HF [24]. Resulting edge roughness from this process was estimated to be less than 300 Å.

8.2.4 Applications Involving Directional Couplers

The dual-channel directional couplers that have been described in this chapter are among the more useful of integrated optic devices. They can be used as power dividers, input or output couplers, and directionally selective taps on an optical data bus. The foregoing are all examples of passive applications, in that the fraction of power coupled in each case is a constant. However, the most important application of the dual-channel coupler is an active modulator or switch, in which the coupled power is electrically controlled. It has been shown in (8.16) and (8.17) that the coupled power is a strong function of any phase mismatch in the two guides. In the next chapter, techniques will be described for producing a change in index of refraction by application of an electric field. The combination of this electro-optic effect with the dual-channel coupler yields a very efficient high-speed modulator, which is described in Chapter 9.

8.3 Butt-Coupled Ridge Waveguides

Techniques of butt-coupling similar to those described for fiber-waveguide couplers in Chapter 7 can also be used to couple two ridge waveguides. Once again, the coupling efficiency depends on the overlap integral of the modes in the waveguides. Koshiba et al. [25] have used the scalar finite element method to evaluate the effective index, modal field profile, and far-field pattern of the guided mode of a ridge waveguide, and to determine the coupling efficiency of a butt-joint structure of two single-mode rib waveguides. Two ridge waveguides of identical cross-section theoretically can have 100% coupling efficiency if they are perfectly aligned and interface reflections are eliminated.

The alignment of coupled ridge waveguides is a critical problem. In the laboratory it can accurately be accomplished by using stages positioned by piezoelectric micrometers, but this is not a satisfactory solution for application in the field. In that case, accurately formed slots or grooves are frequently used to produce the desired alignment. For example, Booth [26] has reported polymer waveguide to waveguide interconnects created by employing laser ablation to micromachine interlocking slots in the waveguides and substrates. Both single-mode and multi-mode coupling have been achieved, with single-mode loss less than 0.2 dB and multi-mode loss of 0.5 dB. The loss has been attributed mostly to overlap mismatch of the abutting guides. Multilayer packaging structures, typically on the order of 200 μm or more in thickness, can be used to increase the strength of the slotted connectors to permit connect/disconnect cycling. Greater than 100 cycles have been tested without any measurable increase in loss.

8.4 Branching Waveguide Couplers

The branching waveguide coupler, or confluent coupler as it is sometimes called, is a straight forward passive combining of a multiplicity of waveguides with a single

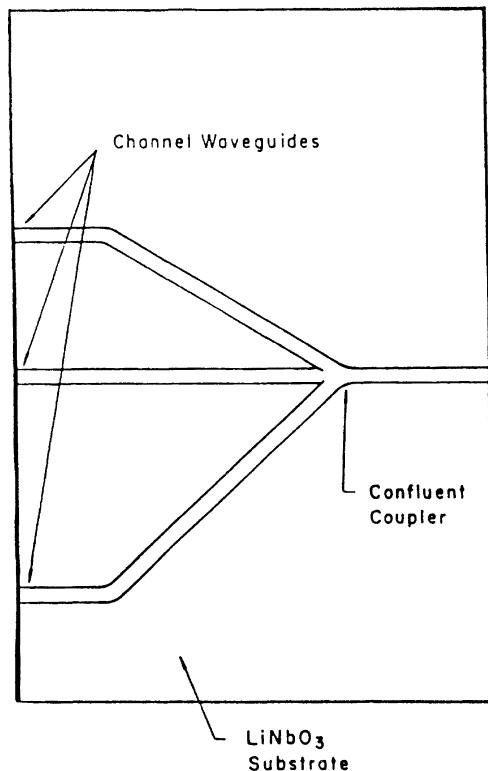


Fig. 8.10 Three-arm branching waveguide coupler

waveguide at a branching junction. It may be used as a combiner, as shown in Fig. 8.10, or as a splitter, if the direction of propagation is reversed from that shown in the figure. The dimensions and refractive index profiles in the waveguides must be carefully controlled in order to obtain either equal splitting or uniform combining. The theory of coupled local normal modes can be used to design branching waveguide couplers. Burns and Milton [27] have applied this approach to both two-branch and three-branch couplers.

Branching waveguide couplers can be produced by any of the usual waveguide fabrication methods. For example, Haruna and Koyama [28] have made them in Ti-diffused Z-cut LiNbO₃ with the input and output waveguide branches 4 μm wide to support only the fundamental guided mode for 6328 Å wavelength. Beguin et al [29] have fabricated branching waveguide couplers by ion exchange in glass. The ion exchange process has two steps. First thallium ions are diffused into the glass at an elevated temperature to cause the exchange of Ti⁺ ions with Na⁺ and K⁺ ions of the substrate, producing an increase in the index of refraction. The guides are then buried and symmetrized by a second ion exchange in the presence of an electric field to prevent out diffusion of the Ti⁺ ions. In 1 \times 2 couplers fabricated by this method,

splitting uniformity was measured to be 0.1 dB and excess loss was less than 0.2 dB. The corresponding values for 1×16 couplers were splitting uniformity = ± 0.3 dB and excess loss < 1.0 dB.

The key feature of branching waveguide couplers which makes them attractive is that they can be relatively inexpensive when produced in large quantities in a low-cost substrate such as glass. For example, they have been proposed for use in the subscriber loops of telecommunication systems [30], where many such devices would be required.

8.5 Optical Fiber Couplers and Splitters

The phenomenon of coupling through the overlapping of evanescent mode tails that has been described for closely spaced rectangular waveguides also provides a means of coupling waves between optical fibers. Dual-channel directional couplers can be conveniently formed from a pair of single-mode optical fibers by clamping the fibers together as shown in Fig. 8.11, heating them until they soften and melt together, then drawing them out to form a fused region of the desired length [31]. The resulting device is called a fused-biconic-taper directional coupler [32, 33].

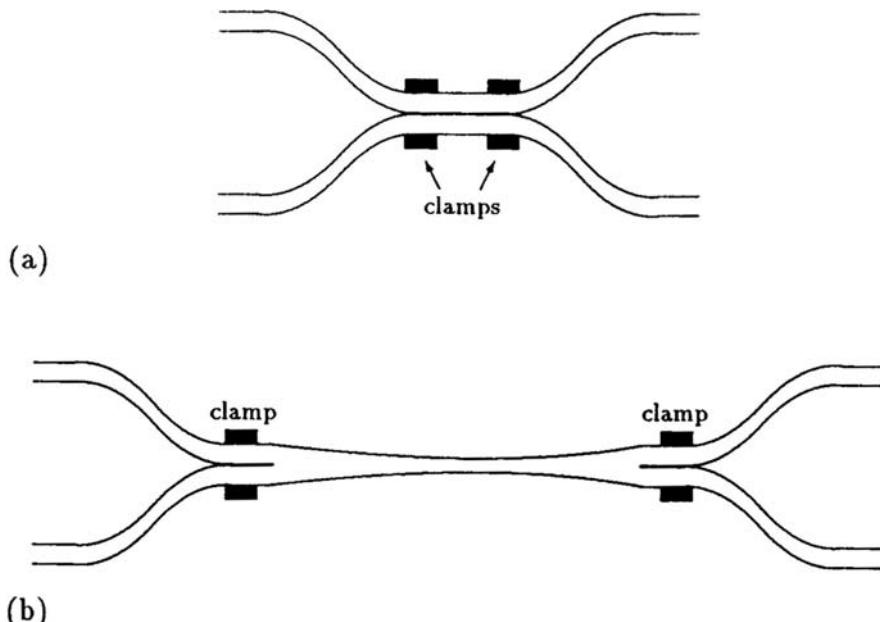


Fig. 8.11 Fabrication method for a fused-biconic-taper directional coupler. (a) Fibers are aligned and clamped next to each other. (b) Fibers are heated and drawn out. Reprinted from [31] p. 134 by Courtesy of Marcel Dekker, Inc.

The coupling can be described by Eqs. (8.1), (8.2), (8.3), (8.4), (8.5), (8.6), (8.7), (8.8), (8.9), (8.10), (8.11), (8.12), (8.13), (8.14) just as was done for rectangular-waveguide directional couplers. By choosing the proper length of the coupling region, the percentage of the power can be adjusted to a desired amount, e.g. to make a 3 dB coupler or a 100% coupler. The magnitude of the coupling coefficient κ depends on the degree of overlap of the mode tails in the coupling region. Since the desired phase-coherent transfer of optical energy requires weak coupling, the cores of the fibers should not overlap. Enough cladding material should remain between them after the fibers are drawn out so that the modes in the two fibers are only weakly coupled. Thus the mode shapes will not be distorted and they will continue to travel with the same velocity, a necessary condition for coherent directional coupling.

The method shown in Fig. 8.11 can also be used to make strongly coupled devices [34, 35]. In that case, the fibers are melted and drawn to a point at which the cores do overlap and fuse together. The coupling action is due to interference between the two lowest order modes of the new waveguide that has formed in the fused region. The starting waveguides can be either single-mode or multimode. The coupling is not coherent as it is in the weakly coupled case. Modes are allowed to mix in the strongly coupled case. However, it is still possible to make splitters with different splitting ratios, depending on the refractive index of the cladding material that is used to encapsulate the coupler and that of the fused cores. More than two fibers can be used in the strongly coupled case to form multi-channel “star” couplers and splitters.

The dual-channel directional couplers and branching waveguide couplers which have been described in this chapter are passive devices that have no controlled switching characteristics. However, by the addition of properly placed control electrodes it is possible to turn these couplers into electro-optic switches and modulators. Devices of this type are described in Chapter 9.

Problems

8.1 A dual-channel-coupler type of modulator has been designed so that

$$\kappa L = \frac{\pi}{2} + m\pi, \quad m = 0, 1, 2, \dots,$$

where κ is the coupling coefficient and L is the length. Thus, complete transfer of light will occur from channel 0 to channel 1. If we now apply a voltage to produce a $\Delta\beta = \beta_0 - \beta_1$, the condition for complete cancellation of the transfer is

$$gL = \pi + m\pi, \quad m = 0, 1, 2, \dots,$$

where

$$g^2 = \kappa^2 + \left(\frac{\Delta\beta}{2} \right)^2.$$

Derive an expression for $\Delta\beta$ required to produce this complete cancellation in terms of the length L .

- 8.2 In order to produce a value of $\Delta\beta$ (as determined in Problem 8.1) that is necessary for complete cancellation, what change in index of refraction in one waveguide (Δn_g) would be required? Given your answer in terms of an expression in which the parameters are wavevectors (k) and coupling length (L).
- 8.3 Using the result of Problem 8.2, what would be the change of index (Δn_g) required for the case of light with a vacuum wavelength $\lambda_0 = 9000 \text{ \AA}$ and a coupling length $L = 1 \text{ cm}$?
- 8.4 If two dual-channel waveguide directional couplers of identical channel geometry and spacing are formed in the same substrate material, except that coupler A has an index of refraction n_A in the channels and coupler B has an index of refraction n_B in the channels, which coupler has the larger coupling coefficient κ if $n_A > n_B$?
- 8.5 A dual-channel directional coupler has $\kappa = 4 \text{ cm}^{-1}$, $\alpha = 0.6 \text{ cm}^{-1}$, and $\Delta\beta = 0$. What length should it be to produce a 3 dB power divider? If that length is doubled, what fraction of the input power is in each channel at the output?
- 8.6 A dual channel directional coupler is made from square $400 \times 400 \text{ nm}$ waveguides with index of refraction = 3.45 formed in a substrate with index = 3.35. The separation between waveguides is 500 nm. For a TE mode with a vacuum wavelength of 900 nm, how long must this coupler be to make a 3 dB coupler if the coupling coefficient $\kappa = 1.229 \times 10^{-5} \text{ nm}^{-1}$?

Hint: You may not need all of the given information to solve this problem.

- 8.7 Explain the mechanism of synchronous coupling that allows 100% transfer of optical power from one waveguide to another in a dual-channel coupler.
- 8.8 In a dual-channel coupler fabricated by masked etching of a planar waveguide layer to produce ridged waveguides on the surface of a lower-index substrate, what can be done to control the magnitude of the coupling coefficient? Include both in-fabrication and post-fabrication methods.

References

1. W.K. Burns, A.F. Milton: Waveguide transitions and junctions, in *Guided-Wave Optoelectronics*, T. Tamir, (ed.), 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) pp. 89–144
2. G.A. Vawter, J.L. Merz, L.A. Coldren: IEEE J. **QE-25**, 154 (1989)
3. K. Utaka, Y. Suematsu, K. Kobayashi, H. Kawanishi: Room-temperature operation of GaInAsP/InP integrated twin-guide lasers with first-order distributed Bragg reflectors. OSA Topical Meeting on Integrated Optics, Incline Village, NV (1980)
4. R. Todt, T. Jacke, R. Meyer, J. Adler, R. Laroy, G. Morthier, M-C. Amann: Sampled grating tunable twin-guide laser diodes with over 40-nm electronic tuning range, IEEE Phot. Technol. Lett. **17**, 2514 (2005)

5. A.J. Baden Fuller: *Microwaves* (Pergamon, Oxford 1979) pp. 237–238
6. S. Somekh, E. Garmire, A. Yariv, R.G. Hunsperger: *Appl. Opt.* **13**, 327 (1974)
7. A. Jervenen, S. Horikanen, S. Najafi: *Opt. Eng.* **32**, 2083 (1993)
8. A. Yariv: *IEEE J. QE*-**9**, 919 (1975)
9. A. Yariv: *Quantum Electronics*, 3rd edn. (Wiley, New York 1989) pp. 623–631
10. R.G. Peall, R.R.A. Syms: *IEEE J. QE*-**7**, 540 (1989)
11. A. von Hippel: Dielectrics, in *Handbook of Physics*. E.U. Condon, H. Odishaw (eds.) 2nd edn., (McGraw-Hill, New York 1967) pp. 4.110–112
12. S. Somekh: Theory, fabrication and performance of some integrated optical devices. PhD Thesis, California Institute of Technology (University Microfilms, Ann Arbor, MI 1974) p. 46
13. H. Garvin, E. Garmire, S. Somekh, H. Stoll, A. Yariv: *Appl. Opt.* **12**, 455 (1973)
14. M.A. Mentzer, R.G. Hunsperger, S. Sriram, J. Bartko, M.S. Wlodowski, J.M. Zavada, H.A. Jenkinson: *Appl. Eng.* **24**, 225 (1985)
15. S. Somekh, E. Garmire, A. Yariv, H. Garvin, R.G. Hunsperger: *Appl. Phys. Lett.* **22**, 46 (1973)
16. E. Garmire, D. Lovelace, G.H.B. Thompson: *Appl. Phys. Lett.* **26**, 329 (1975)
17. N. Schulz, K. Bierwirth, F. Arndt: *IEEE Trans. MTT*-**38**, 722 (1990)
18. W.E. Martin, D.B. Hall: *Appl. Phys. Lett.* **21**, 325 (1972)
19. J. Kondo, K. Aoki, T. Ejiri, Y. Iwata, A. Hamajimal, O. Mitomi, M. Minakata: Ti-diffused optical waveguide with thin LiNbO₃ structure for high-speed and low-drive-voltage modulator, *IEICE Trans. Commun.* **E89-B**, 3428 (2006)
20. L. Riviere, A. Carenco, A. Yi-Yan, R. Guglielmi: Normalized diagrams for diffused waveguides optical properties: Applications to Ti:LiNbO₃ electrooptic directional coupler design, in *Integrated Optics*, H.P. Nolting, R. Ulrich (eds.), Springer Ser. Opt. Sci., Vol. 48 (Springer, Berlin, Heidelberg 1985) pp. 53–57
21. J.C. Campbell, F.A. Blum, D.W. Shaw, K.L. Lawley: *Appl. Phys. Lett.* **27**, 202 (1975)
22. I.P. Kaminov, V. Ramaswamy, R.V. Schmidt, H. Turner: *Appl. Phys. Lett.* **24**, 622 (1974)
23. M. Kawabe, S. Hirata, S. Namba: *IEEE Trans. CAS*-**26**, 1109 (1978)
24. M. Kawabe, M. Kubota, K. Masuda, S. Namba: *J. Vac. Sci. Technol.* **15**, 1096 (1978)
25. M. Koshiba, H. Saitoh, M. Eguchi, K. Hirayama: *IEE Part J. Optoelectron.* **139**, 166 (1992)
26. B.L. Booth: Optical interconnection polymers, in *Polymers for Lightwave and Integrated Optics: Technology and Applications*, A. Hornak (ed.) (Dekker, New York 1992) p. 291
27. W.K. Burns, A.F. Milton: Waveguide transitions and junctions, in *Guided-Wave Optoelectronics*, T. Tamir, (ed.), 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) Chap. 3, in particular, pp. 102–125
28. M. Haruna, J. Koyama: *IEEE J. LT*-**1**, 223 (1983)
29. A. Beguin, T. Dumas, M.J. Hackert: *IEEE J. LT*-**6**, 1483 (1988)
30. P.B. Keck, A.J. Morrow, D.A. Nolan, D.A. Thompson: *IEEE J. LT*-**7**, 1623 (1989)
31. M. Donhowe: Optical fiber waveguides and couplers, in *Photonic Devices and Systems*, R.G. Hunsperger (ed.), (Marcel Dekker, New York, 1994)
32. R.B. Dyott, C.S. Peter, G.A. Clark: Process for optical waveguide coupler, U.S. Patent 3,579,316 9\, May 18,1971)
33. B.S. Kawasaki, K.O. Hill, R.G. Lamont: Biconical-taper single-mode fiber coupler, *Opt. Lett.* **6**, 327 (1981)
34. J. Bures, S. Lacroix, J. Lapierre: Analyse d'un coupleur bidirectionnel à fibres optiques monomode fusionnées, *Appl. Opt.* **22**, 1918 (1983)
35. F.P. Payne, C.D. Hussey, M.S. Yataki: Modelling fused single-mode fibre couplers *Electron. Lett.* **21**, 461 (1985)

Chapter 9

Electro-Optic Modulators

This Chapter begins the discussion of optical-signal modulation and switching. In many cases, the same device can function as either a modulator or a switch depending on the strength of the interaction between the optical waves and the controlling electrical signal, as well as on the arrangement of input and output ports. The device is considered to be a modulator if its primary function is to impress information on a light wave by temporally varying one of its properties. A switch, on the other hand, changes the spatial position of the light, or else turns it off and on. Many of the same factors must be considered in designing or evaluating both modulators and switches. Hence, it is logical to discuss them together.

9.1 Basic Operating Characteristics of Switches and Modulators

9.1.1 Modulation Depth

One important characteristic of modulators and switches is the modulation depth, or modulation index, η . In the case of an intensity modulator in which the applied electrical signal acts to decrease the intensity of the transmitted light, η is given by

$$\eta = (I_0 - I)/I_0, \quad (9.1)$$

where I is the transmitted intensity and I_0 is the value of I with no electrical signal applied. If the applied electrical signal acts to increase the transmitted light intensity, η is given by

$$\eta = (I - I_0)/I_m, \quad (9.2)$$

where I_m is the transmitted intensity when maximum signal is applied. The maximum modulation depth, or extinction ratio, is given by

$$\eta_{\max} = (I_0 - I_m)/I_0 \text{ for } I_m \leq I_0, \quad (9.3)$$

or by

$$\eta_{\max} = (I_m - I_0)/I_m \text{ for } I_m \geq I_0. \quad (9.4)$$

It is also possible to define the modulation depth for phase modulators, as long as the phase change can be functionally related to an equivalent intensity change. For the case of interference modulators, it can be shown that the modulation depth is given by [1,2]

$$\eta = \sin^2(\Delta\varphi/2). \quad (9.5)$$

where $\Delta\varphi$ is the phase change.

Modulation depth has been defined for intensity modulators (and indirectly for phase modulators); however, an analogous figure of merit, the maximum deviation of a frequency modulator, is given by

$$D_{\max} = |f_m - f_0|/f_0 \quad (9.6)$$

where f_0 is the optical carrier frequency, and f_m is the shifted optical frequency when the maximum electrical signal is applied.

9.1.2 Bandwidth

Another important characteristic of modulators and switches is the bandwidth, or range of modulation frequencies over which the device can be operated. By convention, that bandwidth of a modulator is usually taken as the difference between the upper and lower frequencies at which the modulation depth falls to 50% of its maximum value. In the case of a switch, frequency response is usually given in terms of the switching speed, or switching time. The switching time T is related to the bandwidth Δf by the expression

$$T = 2\pi/\Delta f. \quad (9.7)$$

Minimizing switching time is most important when large-scale arrays of switches are used to route optical waves over desired paths. Similarly, modulation bandwidth is a critical factor when many information channels are to be multiplexed onto the same optical beam. Thus, the unusually fast switching speed and wide bandwidth of waveguide switches and modulators, which will be discussed later in this chapter, make them particularly useful in large telecommunications systems.

9.1.3 Insertion Loss

Insertion loss is another important characteristic of optical switches and modulators that must be known for system design. Insertion loss is generally stated in decibels,

and for the case in which the modulating signal acts to decrease the intensity, it is given by

$$\mathcal{L}_i = 10\log(I_t/I_0), \quad (9.8)$$

where I_t is the optical intensity that would be transmitted by the waveguide if the modulator were absent, and I_0 is the intensity transmitted with the modulator in place, but with no applied signal. For a modulator in which the applied signal acts to increase the transmitted intensity, the insertion loss is given by

$$\mathcal{L}_i = 10\log(I_t/I_m), \quad (9.9)$$

where I_m is the transmitted intensity when maximum signal is applied. Insertion loss is, of course, an optical power loss. However, it ultimately increases the amount of electrical power that must be supplied to the system, since higher power optical sources must be used.

9.1.4 Power Consumption

Electrical power must also be supplied to drive the modulator or switch. In the case of modulators, the required drive power increases with modulation frequency. Hence, a useful figure of merit is the drive power per unit bandwidth, $P/\Delta f$, usually expressed in milliwatts per megahertz. As is discussed in more detail in Section 9.7, a key advantage of channel-waveguide modulators is that they have a significantly lower $P/\Delta f$ than that required for bulk modulators.

The power requirements of optical switches operating at high clock rates, for example to time-division multiplex a number of different signals, can be evaluated in much the same way as is used for modulators. Hence, $P/\Delta f$ would still be a useful figure of merit in that case. However, if switching is done at relatively slow rates, a more important quantity is the amount of power required to hold the switch in a given state. An ideal switch would consume significant power only during the change of state; holding power would be negligible. Since electro-optic switches require the presence of an electric field to maintain at least one state, they could not be called ideal in that respect. However, except for leakage current, little power is needed to maintain a field in the small volume of a waveguide switch.

9.1.5 Isolation

The degree of isolation between various inputs and outputs of a switch or modulator is a major design consideration. In a modulator, the isolation between input and output is merely the maximum modulation index, as defined previously. However, it is usually expressed in decibels when used to specify isolation. In the case of a switch, the isolation between two ports (either input or output) is given by

$$\text{isolation [dB]} = 10 \log \frac{I_2}{I_1}, \quad (9.10)$$

where I_1 is the optical intensity in the driving port, and I_2 is the intensity at the driven port when the switch is in the *off* state with respect to port 1 and 2. Thus, a switch with a signal leakage, or crosstalk, of 1% with respect to two ports would have -20 dB isolation.

9.2 The Electro-Optic Effect

The fundamental phenomenon that accounts for the operation of most electrooptic modulators and switches is the change in index of refraction produced by the application of an electric field. In the most general case, this effect is nonisotropic, and contains both linear (Pockels effect) and nonlinear (Kerr effect) components. In crystalline solids, the change in index produced by the linear electro-optic effect, can be most conveniently characterized by the change in the components of the optical indicatrix matrix [3]. The equation of the index ellipsoid in the presence of an electric field is

$$\left(\frac{1}{n^2}\right)_1 x^2 + \left(\frac{1}{n^2}\right)_2 y^2 + \left(\frac{1}{n^2}\right)_3 z^2 + 2 \left(\frac{1}{n^2}\right)_4 yz + 2 \left(\frac{1}{n^2}\right)_5 xz + 2 \left(\frac{1}{n^2}\right)_6 xy = 1 \quad (9.11)$$

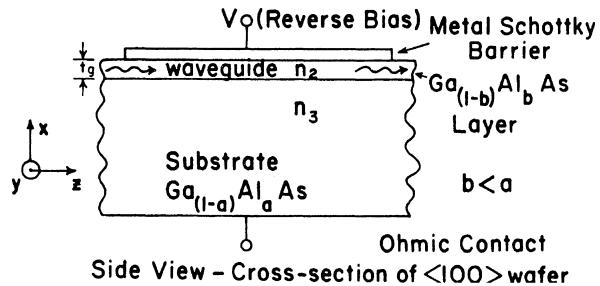
If x , y , and z are chosen to be parallel to the principal dielectric axes of the crystal, the linear change in the coefficients due to an electric field E is given by

$$\Delta \left(\frac{1}{n^2}\right)_i = \sum_{j=1}^3 r_{ij} E_j, \quad (9.12)$$

where $i = 1, 2, 3, 4, 5, 6$, and where $j = 1, 2, 3$ are associated with x, y, z , respectively. If (9.12) is written in matrix form, the $6 \times 3 [r_{ij}]$ matrix is called the electro-optic tensor. It can be shown that the linear electro-optic effect exists only in crystals that do not possess inversion symmetry [4,5]. Even in the case of these noncentrosymmetric crystals, for most symmetry classes, only a few elements of the electro-optic tensor are nonzero [3]. Hence, in the design of an electro-optic modulator or switch, both the waveguide material and its orientation with respect to the applied electric field must be chosen carefully. Nevertheless, many materials that can be used to make low loss waveguides, such as GaAs, GaP, LiNbO₃, LiTaO₃ and quartz, also have significantly large Pockels coefficients, for certain orientation. Thus, the linear electro-optic effect is widely used in integrated optic applications.

The nonlinear (quadratic) Kerr electro-optic coefficient is relatively weak in commonly used waveguide materials. Also, a nonlinear dependence on electric field introduces unwanted modulation crossproducts (distortion) into the modulated signal. Consequently, it is not particularly useful in most integrated optic applications.

Fig. 9.1 Basic electro-optic modulator structure



9.3 Single-Waveguide Electro-Optic Modulators

There are several different types of electro-optic modulators and switches that can be fabricated in a single-waveguide structure. The waveguide may be either a planar, or a channel guide. For example, the relatively simple planar waveguide shown in Fig. 9.1 can function as either a phase modulator, an amplitude (intensity) modulator, or as an optical switch. The waveguide shown in Fig. 9.1 is formed in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ using the heteroepitaxial growth technique described in Chapter 4. However, any electro-optic semiconductor such as $\text{GaAs}_x\text{P}_{(1-x)}$, GaAs , or GaP could be used, and also the waveguide could alternatively be formed by the carrier-concentration-reduction method.

9.3.1 Phase Modulation

The waveguide of Fig. 9.1 is asymmetric, since the indices of the air and the metal at the top surface are both much less than n_2 , while n_3 is relatively close to n_2 . The total change in index between the guide and the substrate is given by

$$\Delta n_{23} = n_2 - n_3 = \Delta n_{\text{chemical}} + \Delta n_{\text{CCR}} + \Delta n_{\text{EO}}, \quad (9.13)$$

where $\Delta n_{\text{chemical}}$ is the index change due to the differing aluminium concentrations a and b , Δn_{CCR} is the change resulting from decreased carrier concentration (if any) in the guide, and Δn_{EO} is the index change caused by the electro-optic effect. To make a phase modulator, the dimensions and doping of the waveguide are chosen so that it is above cutoff for the $m = 0$ mode at the wavelength desired, but is below cutoff for the $m = 1$ mode. Thus, from (3.31)

$$\frac{1}{32n_2} \left(\frac{\lambda_0}{t_g} \right)^2 < \Delta n_{\text{chemical}} + \Delta n_{\text{CCR}} < \frac{9}{32n_2} \left(\frac{\lambda_0}{t_g} \right)^2. \quad (9.14)$$

When a voltage V is applied with reverse bias polarity to the Schottky barrier diode, as shown in Fig. 9.1, the waveguide becomes part of the depletion layer of

the diode, and the electric field causes a change in the phase of light waves traveling along the guide that is proportional to V . For the crystal orientation shown, the change in index of refraction caused by the field for a TE wave (polarized along the y axis) is given by [6]

$$\Delta n_{E0} = n_2^3 r_{41} \frac{V}{2t_g}, \quad (9.15)$$

while there is no field-induced index change for TM waves (polarized in the x direction). By definition of terms,

$$\Delta n = \Delta\beta/k = (\Delta\beta\lambda_0)/2\pi \quad (9.16)$$

Hence, if (9.3.4) is substituted into (9.15), the phase change produced by the electric field is given by

$$\Delta\varphi_{EO} = \Delta\beta L = \frac{\pi}{\lambda_0} n_2^3 r_{41} \frac{VL}{t_g}, \quad (9.17)$$

where L is the length of the modulator in the z direction. The above equations assume that both the optical and electrical fields are uniform and occupy the same volume. If that is not the case, a correction factor Γ can be added to allow for an overlap integral which is less than 1 [6].

Many variations of the basic single-waveguide phase modulator structure have been demonstrated. For example, Hall et al. [7] have made a planar modulator structure like that of Fig. 9.1 in GaAs, in which $\Delta n_{chemical}$ was zero but doping concentration changes in the layers produced a Δn_{CCR} sufficient for waveguiding. Kaminow et al. [8] have made a modulator of this type in an outdiffused LiNbO₃ waveguide structure that was shaped by ion beam etching to produce a 19 μm wide ridged channel waveguide modulator. The required modulating power was only 20 $\mu\text{W}/\text{MHz}/\text{rad}$. In that case, the electric field was introduced into the waveguide by means of two evaporated metal strip electrodes on either side of the waveguide. A voltage of 1.2 V produced a phase change of 1 rad. A reduction of required power to 1.7 $\mu\text{W}/\text{MHz}/\text{rad}$ was obtained in a 5 μm wide Ti-diffused LiNbO₃ single-channel waveguide modulator [9].

Electro-optic modulators can also be made in polymer materials. Shuto et al. [10] have reported electro-optic light modulation and second-harmonic generation in diazo-dye-substituted poled polymers. A poled dimethyl-substituted diazo dye (3RDCVXY) film exhibited a linear EO coefficient of 40 pm/V at 0.633 μm wavelength. An EO channel-waveguide modulator was made in this film with a half-wave voltage of 5 V.

As processing technology has improved, electro-optic modulator research and development has extended into new materials and smaller dimensions. For example, Brosi et al. [11] have designed a novel electro-optic silicon-based modulator with a bandwidth of 78 GHz, a drive voltage amplitude of 1 V and a length of only

80 nm. The interaction region of this device consists of a narrow polymer-filled slot in the center of a photonic crystal waveguide. The design and fabrication of photonic crystals is explained in Chapter 22.

9.3.2 *Polarization Modulation*

Phase modulators are limited in their usefulness by the fact that a phase coherent detection system must be used. To avoid this complication, the simple modulator structure of Fig. 9.1 can be used in a slightly different fashion by introducing the linearly polarized optical beam at 45° to the x and y axes. Because the phase change occurs only for waves polarized in the y direction and not for those polarized in the x direction, a rotation of the polarization vector will result as the waves propagate in the z direction. This change of polarization can be detected with a polarization sensitive detector, or by placing a polarization-selective filter (usually called an analyzer) ahead of the detector. In the case of a discrete waveguide modulator, used for an air beam, a conventional wiregrid polarizer or an absorptive polarizing filter can be used as the analyzer. An analogous system can be implemented for the optical integrated circuit. For example, both grating and prism couplers are polarization sensitive and, hence, can be used as analyzers. However, the difficulty of fabricating an effective analyzer monolithically has limited the use of polarization modulators in OIC's, and has led to a preference for intensity modulation.

9.3.3 *Intensity Modulation*

Since polarization and phase modulation are difficult to detect compared to intensity modulation, the device of Fig. 9.1 is most often used in an intensity modulation mode. To make an intensity modulator, the difference in index of refraction at the waveguide-substrate interface must be carefully tailored to make the waveguide be just at the threshold for guiding the lowest order mode with no electric field present. Then, when an electric field is created by applying a voltage to the electrodes, it produces a slight additional change in index which causes the waveguide to become transmissive. The total change of index is given by (9.13). Hence, the zero field threshold condition for an intensity modulator of this type is given by

$$\Delta n_{23} = \Delta n_{\text{chemical}} + \Delta n_{\text{CCR}} = \frac{1}{32n_2} \left(\frac{\lambda_0}{t_g} \right)^2, \quad (9.18)$$

where (3.31) has been used for the cutoff condition of an asymmetric guide. An intensity modulator of this type was first implemented by Hall et al. [7], in a GaAs carrier-concentration-reduction planar waveguide, for light of $1.15 \mu\text{m}$ wavelength. A voltage of 130 V was required to bring the guide above cutoff for the TE_0 mode. Campbell et al. [12] have made a channel-waveguide modulator of this type

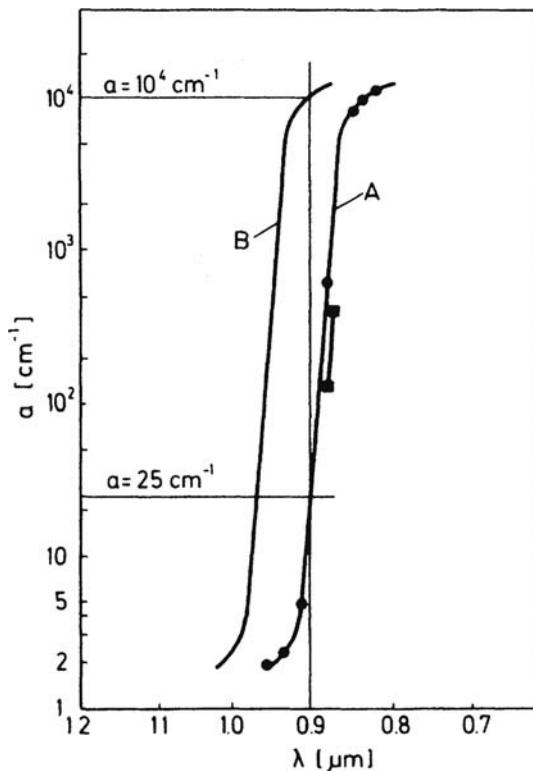
in GaAs, with a 95% maximum modulation depth, a 150 MHz bandwidth, and a power requirement of less than 300 $\mu\text{W}/\text{MHz}$. Kawabe et al. [13] have made an intensity modulator for 6328 Å light in a 2.4 μm wide Ti-diffused LiNbO₃ channel waveguide. With an applied voltage of -10 V (15 kV/cm) the E_{11}^z mode was clearly guided, and with an applied voltage of $+10$ V it was cutoff. The extinction ratio between the two extremes was -19 dB. Obviously, intensity modulators such as have been described in this section, can also function as effective optical switches, as long as the magnitude of ($\Delta n_{\text{chemical}} + \Delta n_{\text{CCR}}$) is chosen to place the waveguide just below threshold and the applied field V/t_g is large enough to bring the guide well above cutoff at the desired wavelength.

9.3.4 *Electro-Absorption Modulation*

The modulators discussed thus far have all depended on the linear electro-optic effect for their operation. However, there is another type of modulator, the electro-absorption modulator, which must be classed as an electro-optic modulator in that it uses an electric field to produce intensity modulation, yet it does not employ the Pockels effect. The electro-absorption modulator is, instead, based either on the Franz-Keldysh effect [14] or the Quantum Confirmed Stark Effect (QCSE). The QCSE is a quantum well property that is covered in Chapter 18. In the following paragraphs of this chapter the discussion will be limited to the Franz-Keldysh type of electroabsorption modulator. In the presence of a strong electric field, the absorption edge of a semiconductor is shifted to longer wavelength, as shown in Fig. 9.2 for the case of GaAs in a field of 1.3×10^5 V/cm. Because of the steepness of the absorption edge in a direct bandgap material such as GaAs, very large changes in absorption of wavelength near the band edge can be produced by application of an electric field. In the specific example shown in Fig. 9.2, for light of 9000 Å wavelength, the absorption coefficient α increases from 25 cm^{-1} to 10^4 cm^{-1} when the field is applied.

The mechanism responsible for the Franz-Keldysh effect can be described straightforwardly with reference to the semiconductor energy band diagram shown in Fig. 9.3. In the presence of a strong electric field, the band edges bend. The left-hand limit of the diagram represents the surface of the semiconductor, at which a Schottky barrier contact or shallow p-n junction has been formed. Application of a reverse-bias voltage to this rectifying junction causes a charge depletion layer to form, extending to depths x within the semiconductor. A nonuniform electric field is present within the depletion layer, with the largest amplitude existing at the surface. The resulting band bending is also greatest at the surface. (For a detailed discussion of depletion layers and band bending see, for example, [15].) Outside of the depletion layer where the field exists, the bands are flat, as at the right side of Fig. 9.3. In this region, a photon can be absorbed only if it has enough energy to lift an electron across the bandgap, as in transition (a). Closer to the surface, where the bands have been bent by the field, a transition (b) can occur in which photon energy is sufficient

Fig. 9.2 Franz-Keldysh shift of the absorption edge of GaAs. *Curve A* is the zero field absorption curve for GaAs. (Circular dots represent experimental data points for n-type material with carrier concentration $n = 3 \times 10^{16} \text{ cm}^{-3}$. For the square dots $n = 5.3 \times 10^6 \text{ cm}^{-3}$.) *Curve B* shows the shifted absorption edge for a field of $1.3 \times 10^5 \text{ V/cm}$



only to lift the electron partway across the gap. Ordinarily such a transition could not occur because there would be no allowed electron state within the bandgap. However, the electric field effectively broadens the states of the conduction band so that there is a finite probability of finding the electron in the gap. This, of course, reduces

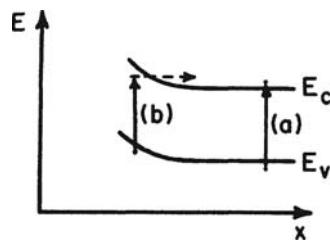


Fig. 9.3 Energy band diagram of a semi-conductor exhibiting the Franz-Keldysh effect in the presence of a strong electric field. The parameter x represents the distance from the surface of the semiconductor, and E is the electron energy. E_c and E_v are the conduction and valence bandedges, respectively

the effective bandgap and thereby shifts the absorption edge to longer wavelength. It can be shown that the effective change in bandgap energy ΔE is given by [14, 16]

$$\Delta E = \frac{3}{2} (m^*)^{-1/3} (q\hbar\varepsilon)^{2/3}, \quad (9.19)$$

where m^* is the effective mass of the carrier, q is the magnitude of the electrical charge of a carrier, \hbar is Planck's constant divided by 2π , and ε is the electric field strength.

Since ΔE depends on the electric field strength, and since α is a very strong function of ΔE , as shown in Fig. 9.2, a very effective electro-absorption modulator can be made for light of slightly less than bandgap wavelength. A basic electro-absorption modulator structure is shown in Fig. 9.4. The surface electrode can be either a Schottky barrier contact or a shallow p-n junction. In either case the electric field is produced in the depletion layer. Ideally the dopant concentration in the waveguide N_2 should be small enough that the depletion layer extends all the way through the guide in the x direction. The length of the modulator, and the applied voltage are chosen by using absorption curves such as those of Fig. 9.2 in order to establish a desired minimum insertion loss and maximum modulation depth for a given wavelength.

Although any type of waveguide can be used, the waveguide modulator structure can be improved by using a GaAlAs heterostructure guide, as shown in Fig. 9.4. In that case, the aluminum concentration in the guide can be adjusted to produce optimum performance at a given wavelength. For example, Reinhart [17] has produced such structures for use at 9000 Å which exhibit a change in transmission by a factor of 100 for -8 V applied bias. The power necessary for 90% modulation was on the order of 0.1 mW/MHz. By using multilayer waveguide structures of the quaternary compound GaInAsP, electroabsorption modulators for use at 1.3 or 1.55 μ m wavelength can be made [18, 19].

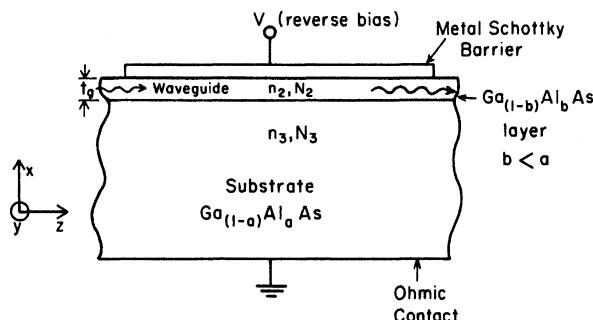


Fig. 9.4 Basic electro-absorption modulator. Aluminum concentration b in the waveguide should be chosen so that the absorption edge wavelength is just slightly shorter than the guided wavelength, i.e., the guide is transparent with $V = 0$. Concentrations should be chosen so the $N_3 >> N_2$, thus, a relatively large electric field is produced in the guide when V is applied

In Section 9.3, electro-optic modulators employing a single waveguide have been discussed. In the next section, we consider modulators involving the transfer of optical energy between two waveguides. Such devices are inherently switches as well as modulators.

9.4 Dual-Channel Waveguide Electro-Optic Modulators

In Chapter 8, we saw that two closely spaced channel waveguides could function as a directional coupler, in which optical energy was synchronously transferred from one guide to the other. Such a coupler can be made into an electro-optic modulator by merely adding two electrodes, as shown in Fig. 9.5.

9.4.1 Theory of Operation

If a modulating signal voltage is applied to the electrodes, as shown in Fig. 9.5, it produces a slight difference in the indices of refraction in the guides, which results in a propagation constant difference $\Delta\beta$. By once again following the coupled mode theory approach used in Chapter 8 [20], the coupling equations can be shown to be given by

$$\frac{dA_0(z)}{dz} = -i\beta_0 A_0(z) - ik A_1(z), \quad (9.20)$$

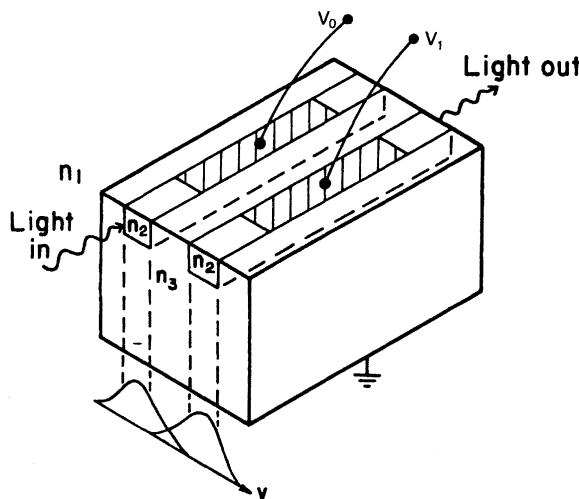


Fig. 9.5 Basic dual-channel modulator structure

and

$$\frac{dA_1(z)}{dz} = -i\beta_1 A_1(z) - ik A_0(z), \quad (9.21)$$

where β_0 and β_1 are the propagation constants in the two guides, and the other terms have already been defined previously. The solution of (9.20) and (9.21), subject to the boundary conditions that

$$A_0(0) = 1 \quad \text{and} \quad A_1(0) = 0, \quad (9.22)$$

yields the following expressions for $A_0(z)$ and $A_1(z)$.

$$A_0(z) = \left(\cos(gz) - i \frac{\Delta\beta}{2g} \sin(gz) \right) \exp \left[-i \left(\beta_0 - \frac{\Delta\beta}{2} \right) z \right], \quad (9.23)$$

and

$$A_1(z) = - \left(\frac{-ik}{g} \sin(gz) \right) \exp \left[-i \left(\beta_1 + \frac{\Delta\beta}{2} \right) z \right], \quad (9.24)$$

where

$$\Delta\beta = \beta_0 - \beta_1$$

and where

$$g^2 \equiv k^2 + \left(\frac{\Delta\beta}{2} \right)^2. \quad (9.25)$$

Thus, in this case of imperfect phase match, the power flow in the two guides is given by

$$P_0(z) = A_0(z)A_0^*(z) = \cos^2(gz)e^{-\alpha z} + \left(\frac{\Delta\beta}{2} \right)^2 \frac{\sin^2(gz)}{g^2} e^{-\alpha z}, \quad (9.26)$$

and

$$P_1(z) = A_1(z)A_1^*(z) = \frac{\kappa^2}{g^2} \sin^2(gz)e^{-\alpha z}, \quad (9.27)$$

where α is the exponential loss coefficient in the guides. Note that (9.26) and (9.27) become identical to (8.12) and (8.13) when $\Delta\beta$ equals zero. Thus, the condition for total transfer of power for zero applied voltage is given once again by (8.14), which states that

$$\kappa L = \frac{\pi}{2} + m\pi. \quad (9.28)$$

where $m = 0, 1, 2, \dots$. Similarly, it can be seen from (9.26) and (9.27) that, when a modulating voltage is applied to produce a $\Delta\beta$, the coupling is completely cancelled [i.e., $P_1(L) = 0$ and $P_0(L) = 1$] if

$$gL = \pi + m\pi, \quad \text{where } m = 0, 1, 2, \dots \quad (9.29)$$

From (9.28) and (9.29) it can be shown that the value of $\Delta\beta$ required for 100% modulation is given by

$$(\Delta\beta)L = \sqrt{3}\pi. \quad (9.30)$$

The effective index of refraction in a guide is given by

$$n_g \equiv \frac{\beta}{k}. \quad (9.31)$$

Thus, the change in effective index needed for 100% modulation is given by

$$\Delta n_g = \frac{\sqrt{3}\pi}{kL}. \quad (9.32)$$

In typical cases, the magnitude of Δn_g required for 100% modulation is surprisingly small. For example, in a GaAlAs $3 \mu\text{m} \times 3 \mu\text{m}$ dual-channel waveguide modulator such as that of Fig. 9.5, of length equal to 1 cm, (9.32) predicts that light of 9000 Å vacuum wavelength can be totally switched by producing a Δn_g of approximately 1×10^{-4} . From (9.15) it can be determined that the magnitude of the required electric field is about 3×10^4 V/cm, corresponding to a voltage of 10 V across the $3 \mu\text{m}$ thick channels.

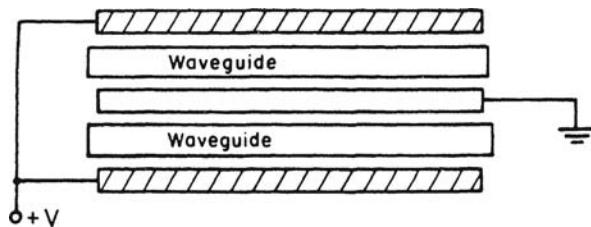
The particular modulator geometry shown in Fig. 9.5 is probably the simplest arrangement that can be envisioned, so that it serves as a good example to illustrate the principles of dual-channel modulator operation.

However, many other electrode configurations are possible, some of them having distinct advantages for given applications.

9.4.2 Operating Characteristics of Dual-Channel Modulators

The concept of using a dual-channel directional coupler as a modulator was proposed as early as 1969 by Marcatili [21], but quite a few years passed before a working device was realized, largely because of the difficulty of fabricating the dual-channel structure to the required close tolerances. Somekh et al. [22,23] fabricated dual-channel directional couplers in GaAs with 100% coupling, and theoretically analyzed the case of non-zero $\Delta\beta$. Taylor [24] theoretically analyzed the performance of a dual-channel coupler with three electrodes, as shown in Fig. 9.6; and Campbell et al. [25] produced the first operational device of this type, in 1975. At

Fig. 9.6 Topographic view of dual-channel modulator with three electrodes



about the same time Papuchon et al. [26,27] reported the successful operation of the first two-electrode type modulator, such as that shown in Fig. 9.5.

The three-electrode device of Campbell et al. was fabricated in GaAs, by using the metal electrodes themselves to create a pair of strip-loaded waveguides, as shown in Fig. 9.7. Ninety-five percent amplitude switching, with a maximum extinction ratio of 13dB, was observed for the light from a 1.06 μm Nd:YAG laser. For

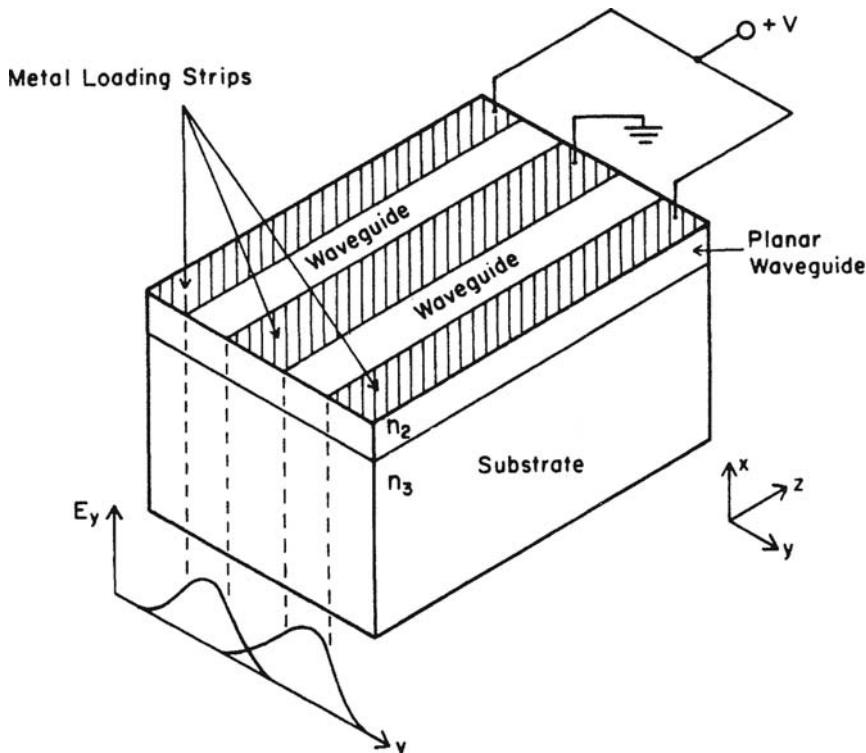


Fig. 9.7 Dual-channel modulator featuring strip loaded waveguide [23]. The three loading strips also function as the Schottky Barrier electrodes of the modulator. Typical electric field distribution for the TE_0 modes are plotted beneath the modulator structure

the case of $6 \mu\text{m}$ wide waveguides separated by $7 \mu\text{m}$, the maximum switching condition occurred for an applied voltages of 35 V . A 7 ns rise-time was measured, implying a 3 dB bandwidth of 100 MHz . The power-bandwidth ratio was determined to be approximately $180 \mu\text{W/MHz}$.

The frequency response of modulators and switches of this type is generally limited by the capacitance of the electrodes. For example, Campbell et al. estimate that the above mentioned 7 ns rise-time could be reduced by an order of magnitude by reducing the width of the Schottky barrier contacts from 100 to $10 \mu\text{m}$, thus giving a projected bandwidth of 1 GHz . The two-electrode modulator of Papuchon et al [26,27] offers the potential advantage of lower capacitance by eliminating the center electrode. They called their modulator *commutateur optique binaire rapide*, leading to the acronym COBRA which is sometimes used to label the two-electrode type of dual-channel modulators. Typical coupling lengths observed for couplers formed from $2 \mu\text{m}$ wide Ti diffused strips, in LiNbO_3 , and used to guide 5145 \AA light, were $5000 \mu\text{m}$ for the case of $2 \mu\text{m}$ separation, and 1 mm for $3 \mu\text{m}$ spacing. A voltages of 6 V was found to be sufficient to *switch off* the coupling that was present for 0 V applied. The effect was about three times stronger for TM guided waves than for TE because of the particular orientation of the field and the LiNbO_3 substrate, and the strong anisotropy of the electro-optic tensor.

A problem that is shared by both the two-electrode and three-electrode modulators discussed thus far is that the length of the device must be carefully chosen to establish maximum coupling with no applied bias, as per (9.28); the on state cannot be electrically adjusted. Kogelnik and Schmidt [27] have demonstrated a dual-channel modulator in which three basic electrodes are split in half, as shown in Fig. 9.8. The polarity of the applied voltage is reversed in the two halves to produce two sections with $\Delta\beta$ of equal magnitude, but opposite sign. The effect of this stepped $\Delta\beta$ reversal is to yield a device in which both the *off* and *on* states can be electrically adjusted for a relative wide range of lengths. Obviously this allows one to maximize the extinction ratio and minimize crosstalk. For example, Kogelnik and Schmidt [27] have shown that providing sections with alternating $\Delta\beta$ makes it possible to cause a complete transfer of the light from one waveguide to another by electrical adjustment of the on and off states, as long as the radio L/l is greater than unity, where L is the total length of the modulator, and l is the effective interaction length, defined by $l = \pi/2\kappa$. The frequency response of dual-channel modulators is also improved

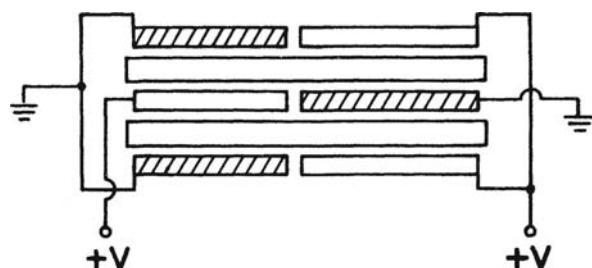


Fig. 9.8 Topographic view of dual-channel modulator with split electrodes to produce a stepped $\Delta\beta$ reversal [27]

by stepped $\Delta\beta$ reversal [28]. For a detailed discussion of the theory of reversed $\Delta\beta$ modulators see Alferness [29]. During the years following its inception the effectiveness of the stepped $\Delta\beta$ reversal modulator has been greatly improved by the use of more sophisticated device structures. For example, Veselka et al. [30] have produced a 1×4 switch array of these devices (intended for time-division-multiplexing applications) which switches at a frequency of 4 GHz through the use of traveling wave electrodes. The high-frequency advantages of traveling wave electrodes are discussed in Section 9.8.

Because of the anisotropic nature of the electro-optic tensor, electro-optic modulators are generally sensitive to the polarization of the light waves. In many applications this is not a problem because the proper polarization can be chosen to maximize the desired interaction. However, in the case of modulators to be used in conjunction with fiber optic waveguides, polarization sensitivity becomes an important problem to be dealt with. Linearly polarized light coupled into circular, single-mode fibers undergoes conversion to elliptical polarization, which excites both TE and TM waves in the rectangular guide of an OIC. Thus, in order to obtain maximum extinction ratio and minimum crosstalk, switches and modulators must be carefully designed to act on both polarization components equally. Steinberg and Giallorenzi [31] have shown that this can be done for most types of electro-optic modulators in LiNbO_3 by properly selecting the crystal orientation. Steinberg et al. [32] have also shown that a polarization-insensitive modulator can be made by using the unique electrode configuration of Fig. 9.9, which combines both between-guide electrodes and over-guide electrodes in a stepped $\Delta\beta$ reversal pattern. Since the between-guide electrodes produce electric field lines parallel to the substrate plane, while the over guide electrodes produce lines perpendicular to it, the designer is given an additional degree of freedom that can be used to cancel polarization sensitivity.

Thus far the discussion of dual channel waveguide couplers and modulators has been in regard to devices in which controllable coupling of lightwaves from one waveguide to another has been desired. However, as optical integrated circuits have become more complex and the packing of components has become more dense, the problem of undesired coupling between two adjacent waveguides has become

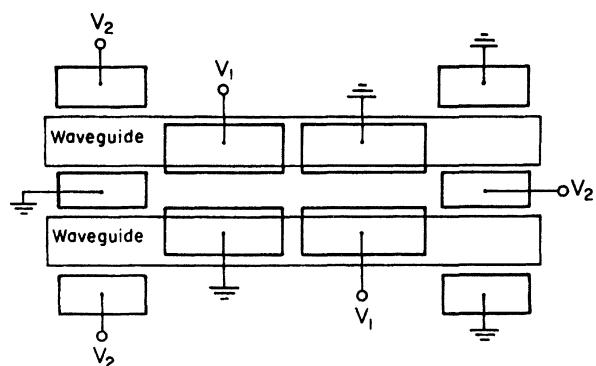


Fig. 9.9 Electrode configuration of a polarization insensitive modulator [32]

one which must be considered during the design phase. Such undesired coupling, or “crosstalk” occurs when the spacing between waveguides is so close that the evanescent mode tails overlap. Crosstalk must be eliminated or at least minimized because it drains energy out of the lightwave signal channel and also introduces an undesired lightwave signal into the adjacent waveguide. Computation of crosstalk between single-mode rectangular optical waveguides for application to planar light-guide circuit design has been considered by Sikorski et al. [33].

9.5 Mach-Zehnder Type Electro-Optic Modulators

Not all two-channel modulators employ the synchronous coupling of energy between overlapping mode tails. Another class of modulators is based on a waveguide version of the **Mach-Zehnder interferometer**, in which interference is produced between phase coherent light waves that have traveled over different path lengths. The basic modulator structure is shown in Fig. 9.10. Light input to the modulator is via a single-mode waveguide. A beam splitter divides the light into two equal beams that travel through guides *a* and *b*, respectively. By applying a voltage to the electrodes, the effective path lengths can be varied. In an ideally designed modulator of this type, the path lengths and guide characteristics are identical, so that with no applied voltage the split beams recombine in the output waveguide to produce the lowest-order mode once more. If an electric field is applied so as to produce a phase change of π radians between the two arms, then the recombination results in an optical field that is zero at the center of the output waveguide, corresponding to the first order ($m = 1$) mode. If the output waveguide is a single mode guide, identical to the input guide, the first order mode is cut off, and rapidly dissipates over a short length by substrate radiation. Thus, the modulator can be switched from a transmitting to a nontransmitting state by application of a voltage.

Various embodiments of the basic Mach-Zehnder interferometer structure of Fig. 9.10 have been demonstrated to be effective. Zernike [34] first proposed this type of modulator in integrated optic form, using 3-dB directional couplers as the

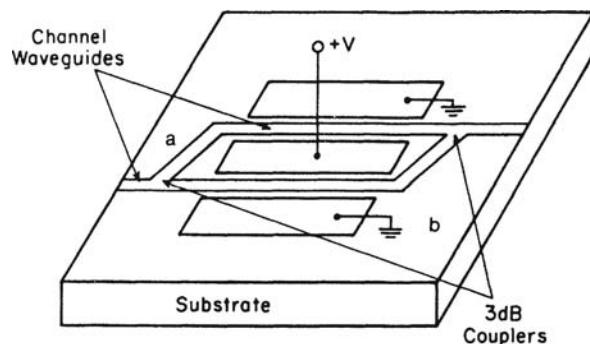


Fig. 9.10 Mach-Zehnder type modulator

beam splitter and recombiner, rather than the prisms used in a conventional discrete component interferometer. Martin [35] fabricated a Mach-Zehnder modulator in ZnSe, employing diffused waveguides that were formed into single mode branching "Y" s to obtain the desired beam splitting and recombining. Operating at a wavelength of 0.63 μm , this device could be switched from a state of 60% transmission to about 1% transmission by application of approximately 25 V. Mach-Zehnder modulators can be fabricated in LiNbO_3 substrates by forming either Ti indiffused or proton exchanged waveguides [36]. They can also be fabricated in spun-on polymer materials [37]. One problem with all of the Mach-Zehnder modulators described thus far is that even minute variations in fabrication parameters result in a device which is not in the *on* state for zero applied voltage. Thus, careful control of the applied voltage must be maintained for both the *off* and *on* states. Ramaswamy et al. [38] have gone one step further by using electrically switched dual-channel directional couplers as the beam splitter and recombiner. This modification permits electrical adjustment of both the splitting fractions and the phase change in the interferometer arms. Thus, the *off* and *on* states can be electrically selected for optimum extinction ratio. Once the bias voltages on the directional couplers have been adjusted, only the voltage on the interferometer arm must be changed to control switching. An off-on ratio of 22 dB was observed in a 38 mm long device of this type formed in Ti diffused LiNbO_3 .

9.6 Electro-Optic Modulators Employing Reflection or Diffraction

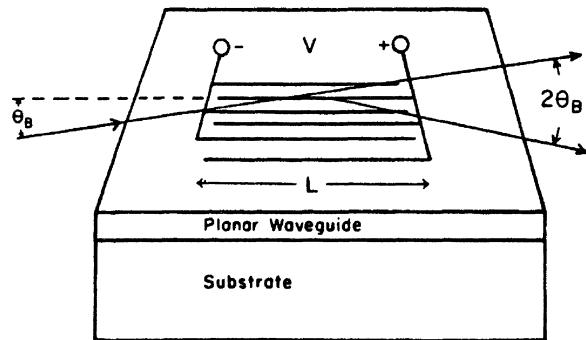
A number of different modulators and switches have been demonstrated that utilize electro-optic control of either reflection or diffraction of the waveguided light. Diffraction modulators are generally based on the Bragg effect [39], which involves distributed interactions with multiple reflecting elements, usually in the form of an optical grating.

9.6.1 Bragg-Effect Electro-Optic Modulators

A typical Bragg-effect modulator is shown in Fig. 9.11, which consists of an interlaced, comb like, pair of electrodes. A voltage applied to the inter-digitated surface electrodes perturbs the index of refraction beneath them thus forming an effective optical grating pattern in the waveguide. This grating causes a change in the direction of propagation of the light beam. If the direction of the light beam in the waveguide is adjusted so it is incident onto the grating bars at an angle equal to the Bragg angle θ_B , the light is diffracted with maximum efficiency at an angle $2\theta_B$ with respect to the input beam. It can be shown [40] that θ_B is given by

$$\sin \theta_B = \lambda_0 / 2\Lambda n_g, \quad (9.33)$$

Fig. 9.11 Bragg-effect electro-optic modulator



where Λ is the grating spacing, and n_g is the effective guide index (β/k).

The derivation of (9.33) is based on the *thick grating* assumption that

$$2\pi\lambda_0 L \gg \Lambda^2. \quad (9.34)$$

If the input beam strikes the grating at an angle different from the Bragg angle, diffraction still occurs over a limited range of $\Delta\theta_B$, but with reduced efficiency. The angular range for a 50% reduction is given by [36]

$$\Delta\theta_B = \frac{2\Lambda}{L}, \quad (9.35)$$

for small $\Delta\theta_B$ such that $\sin\theta_B \approx \theta_B$.

The intensity of light diffracted is dependent on the applied voltage, and has the general form [9.36]

$$\frac{I}{I_0} = \sin^2 VB. \quad (9.36)$$

where I is the intensity of the diffracted beam with V applied, I_0 is the transmitted intensity with $V = 0$, and B is a constant dependent on the effective guide index and on the applicable element of the electro-optic tensor.

Electro-optic Bragg diffraction modulators were first proposed by Hammer [41], and independently by Giarusso and Harris [42]. Hammer et al. demonstrated efficient grating modulation in epitaxial ZnO waveguides on sapphire substrate [43] and in $\text{LiNb}_x\text{Ta}_{(1-x)}\text{O}_3$ waveguides on LiTaO_3 substrates [44]. Detailed theoretical models for this type of modulator have been developed by Lee and Wang [45] and Xin and Tsai [46]. Experimental work [45, 46] has shown that very efficient Bragg diffraction electro-optic modulators can be made if grating geometry and tolerances are carefully controlled. For example, Tangonan et al. [47] formed modulators in Ti diffused LiNbO_3 having a multiple, series grating structure designed for high efficiency. They measured diffraction efficiencies of 98%, with an extinction ratio of

300:1 (24.7 dB) at 1.06 μm wavelength, and a diffraction efficiency of 98%, with an extinction ratio of 250:1 (24 dB) at 6328 Å.

9.6.2 Electro-Optic Reflection Modulators

It is possible to use the linear electro-optic effect to reduce the index of refraction in a layer, thereby bringing about the total internal reflection of an optical beam. A device of this type, which has been proposed by Tsai et al. [48], is shown in Fig. 9.12. Four horn-shaped tapered channel waveguides form the input and output ports to a planar waveguide modulator which contains a region in which the refractive index can be reduced by application of an electric field. If there is not applied voltage, an incident light beam from, for example, port I will encounter no index interface and will pass freely to port 4. If the horn tapers are carefully designed and fabricated to minimize scattering and mode conversion, very little crosstalk will occur at port 3. However, when a voltage is applied with the proper polarity to reduce the index between the electrodes, two interfaces are created between regions of different index. Total internal reflection may occur at the first interface if the angle of incidence is greater than the critical angle, thereby causing partial (or possible total) switching of the light beam to port 3. It can be shown for the arrangement diagrammed in Fig. 9.12 that the critical angle is given by [48]

$$\theta_c = \sin^{-1} \left[1 - \frac{1}{2} n_1^2 r_{33} \left(\frac{V}{d} \right) \right], \quad (9.37)$$

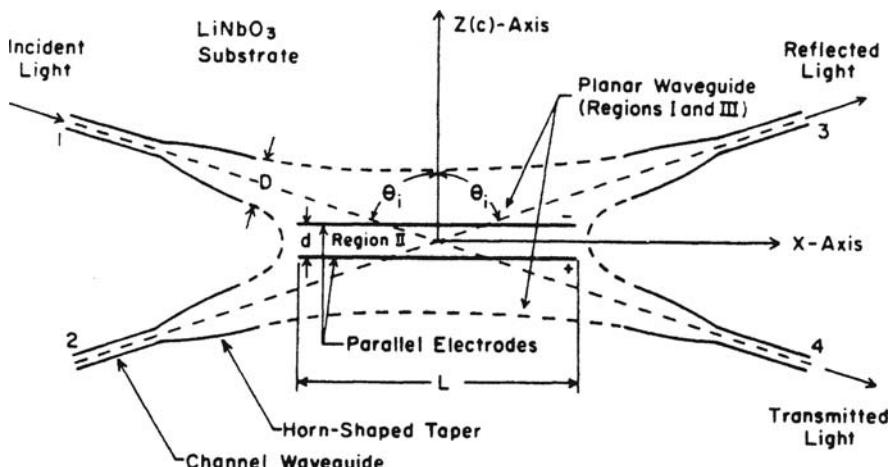


Fig. 9.12 A total internal reflection (TIR) electro-optic modulator and switch [48]

where n_1 is the effective index outside of the electric field region, and d is the electrode separation. In terms of the voltage required to switch a beam incident at a given angle θ_i , (9.6.5) can be written as

$$\left. \frac{V}{d} \right|_{\text{TIR}} = \frac{2(1 - \sin \theta_i)}{n_1^2 r_{33}} \cong \frac{1}{n_1^2 r_{33}} \left(\frac{\pi}{2} - \theta_i \right)^2. \quad (9.38)$$

A TIR switch of the type which has been described has been fabricated by Tsai et al. [48] by Ti diffusion of Y cut LiNbO₃. The input/output horns were 4.7 mm long, tapering from 4 to 40 μm in width. The length of the electrode pair was $L = 3.4$ mm and d equaled 4 μm. Complete switching of a 6328 Å beam was observed for V approximately equal to 50 V. Cross-talk to port 3 in the absence of an applied voltage was measured to be −15 dB. A detailed theoretical model for this type of TIR switch has been developed by Sheem [49]. Estimated switching speeds exceed 6 GHz because of the relatively low device capacitance. Oh et al. [50] have fabricated a TIR optical switch like that of Fig. 9.12 in which the controllable reflecting interface is provided by p/n/p/n current blocking layers. This device, fabricated in InGaAsP waveguides on an InP substrate, had a very low operating current of 20 mA.

9.7 Comparison of Waveguide Modulators to Bulk Electro-Optic Modulators

At several places in this chapter, the relatively low drive power required by a waveguide modulator has been noted. To quantitatively compare the power requirements of a waveguide modulator to those of a bulk electro-optic modulator, it is convenient to develop a simple, yet general, expression for the average external power P_e needed to operate the modulator at a maximum frequency equal to its bandwidth (Δf). For the case of 100% modulation, this power is given by

$$P_e = (\Delta f)\mathcal{W}, \quad (9.39)$$

where \mathcal{W} is the energy supplied from an external source to switch the device on or off. For an ideal electro-optic modulator with no Ohmic losses, all of this energy goes into the stored electric field between the electrodes. Hence, we can take

$$\mathcal{W} = \frac{1}{2} \int \varepsilon E_a^2 dv, \quad (9.40)$$

where E_a is the peak amplitude of the applied field, and ε is the permittivity. The integral is to be taken over the entire volume occupied by the field. If we assume, for convenience, that all of the electric field is confined to the modulator volume, and additionally that E_a is uniform over that volume, (9.40) becomes

$$\mathcal{W} = \frac{\varepsilon W t L E_a^2}{2}, \quad (9.41)$$

where W is the width, t is the thickness and L is the length of the modulator active volume. Hence, the external drive power, from (9.39), is given by

$$P_e = \frac{(\Delta f)\varepsilon W t L E_a^2}{2} \quad (9.42)$$

The key feature of (9.42) is that the modulating power required is proportional to the active volume. Thus, if we compare a bulk electro-optic modulator, such as that shown in Fig 9.13a, to the planar waveguide modulator of Fig. 9.13b, it is obvious that significantly less power is required by the planar waveguide device. Still greater power reduction is obtained by going to a channel waveguide modulator, such as that shown in Fig. 9.13c. Consider the following numerical example.

For the specific case of an electro-optic modulator formed in GaAs with the orientation shown in Fig. 9.1, we find by using (9.15) that the applied field and the resulting change in index of refraction are related by

$$E_a = \frac{2\Delta n}{n_2^3 r_{41}}. \quad (9.43)$$

Therefore, combining (9.43) and (9.42) yields

$$P_e = \frac{2(\Delta f)\varepsilon W t L}{n_2^6 r_{41}^2} \Delta n^2. \quad (9.44)$$

For the special case of the dual-channel 100% modulator, it has been shown in (9.32) that

$$\Delta n = \frac{\sqrt{3}\pi}{kL} = \frac{\sqrt{3}\lambda_0}{2L}. \quad (9.45)$$

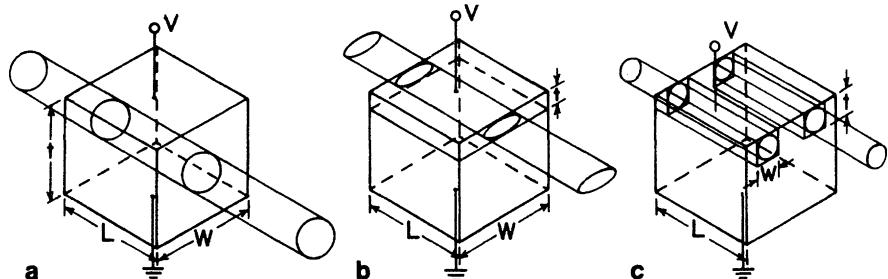


Fig. 9.13 a–c Basic electro-optic modulator structures; **a** bulk; **b** planar waveguide; **c** channel waveguide

Substituting (9.45) into (9.44) gives the expression

$$\frac{P_e}{(\Delta f)} = \frac{3\epsilon W t \lambda_0^2}{2n_2^6 r_{41}^2 L}. \quad (9.46)$$

If we take the following typical values for GaAs: $W = 6 \times 10^{-6}$ m, $t = 3 \times 10^{-6}$ m, $\lambda_0 = 0.9 \times 10^{-6}$ m, $n_2 = 3.6$, $r_{41} = 1.2 \times 10^{-12}$ m/V, $\epsilon/\epsilon_0 = 12$, and $L = 0.5$ cm, the result is that $P_e/(\Delta f)0 = 0.148$ mW/MHz. The comparable value of $P_e/(\Delta f)$ for a planar waveguide modulator would typically be on the order of ten times larger because of the increases in W , while that for a bulk modulator would be 100 to 1000 times larger because of corresponding increases in both W and t . It should be noted that the values of $P_e/(\Delta f)$ calculated using (9.44) are based on the assumption that the optical fields and the electric fields are both uniformly confined to a volume W,L . If this is not the case, a slightly modified relation can be used, which is given by

$$P_e = \frac{2(\Delta f)\epsilon \left(\frac{W}{c_1}\right) \left(\frac{t}{c_2}\right) L}{n_2^6 r_{41}^2} \Delta n^2, \quad (9.47)$$

where c_1 and c_2 are constants, having a values less than 1, to account for electric field and optical field not being perfectly confined to the same volume.

9.8 Traveling Wave Electrode Configurations

The various electro-optic modulators shown in the preceding figures of this chapter all have “lumped-element” electrodes. In this case resistance, capacitance and inductance are all assumed to be non-distributed, and the electrode is considered as an equipotential element. While this model is accurate at lower modulation frequencies, it cannot be used above about 1 GHz. At these higher frequencies the magnitude of the wavelength of the modulating signal approaches the dimensions of the electrode, and such parameters as resistance and capacitance must be considered to be distributed. The modulating signal produces a voltage wave, with peaks and troughs along the length of the electrode. Thus for efficient coupling of the modulating signal the electrodes should be designed so as to from an appropriate microwave transmission line. Traveling wave electrode configurations for a phase modulator and a Mach-Zehnder modulator are shown in Fig. 9.14. In these devices the electrodes have been patterned to form a microwave transmission line with a characteristic impedance of 50Ω . The modulation signal input is at the left end of the transmission line and the modulation signal output at the right end of the line would be terminated in a matching impedance to avoid reflections. As the microwave modulation signal travels along the transmission line the traveling voltage wave produces a similarly moving pattern of changed index of refraction in the underlying waveguide. This change in index causes a phase shift in the optical waves in the waveguide. Because of efficient coupling between the traveling modulating

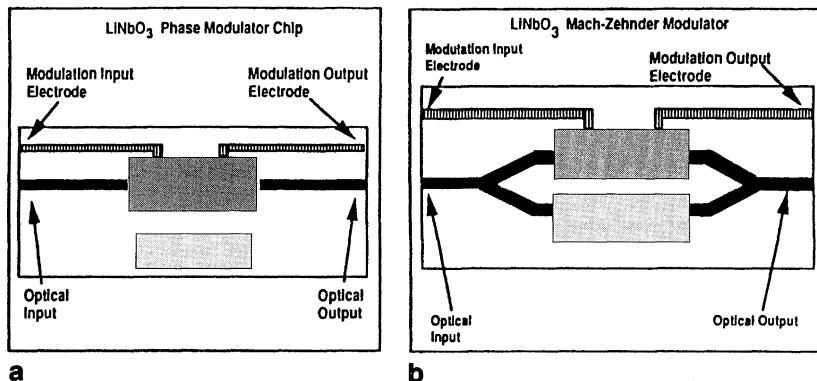


Fig. 9.14 a,b. Waveguide modulators with traveling wave electrodes in LiNbO_3 , **a** phase modulator, **b** Mach-Zehnder modulator

voltage wave and the optical waves the modulators shown in Fig. 9.14 have a 3 dB bandwidth of 4–6 GHz. Similar devices made with conventional electrodes have a bandwidth of only 2–3 GHz.

The characteristic impedance (z) of a traveling wave electrode is given by

$$\frac{1}{z} = \frac{c}{\sqrt{\epsilon_{\text{eff}}}} \left(\frac{C}{L} \right), \quad (9.48)$$

where c is the speed of light in vacuum, (C/L) is the capacitance per unit length and ϵ_{eff} is the effective microwave dielectric constant ($= 35.8$ for lithium niobate) [6]. For a thorough discussion of traveling wave electrode design and performance see Alferness [7].

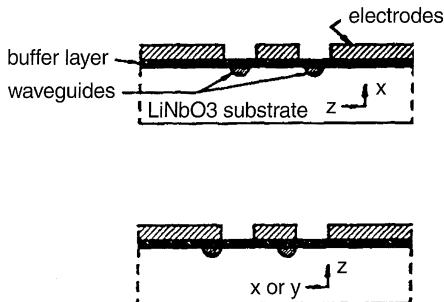
By carefully designing the traveling-wave electrode configuration it is possible to produce electro-optic modulators with very high operating frequencies. The key to success is to match the velocity of the light waves in the optical waveguide to the velocity of the microwaves on the electrode transmission line.

For example, Jungerman and Dolfi [51] have reported a Mach-Zehnder modulator that featured the coplanar waveguide structure shown in Fig. 9.15. The waveguides were made by titanium diffusion into the LiNbO_3 substrate. The use of a buffer layer in conjunction with narrow ground planes and thicker electrodes increased the microwave velocity to match that of the optical signal. The position of the electrodes relative to the waveguides was different for x -cut and z -cut substrates in order to place the electric field of both the microwave and optical signals along the z -axis of the crystal to maximize the electrooptic modulation.

The modulator had an insertion loss < 4 dB and required a 10.4 volt signal for full off-on switching. The modulation 3 dB bandwidth was 50 GHz. The device operated at wavelengths of both 1300 nm and 1550 nm.

Teng [37] has made a Mach-Zehnder modulator with traveling wave electrodes in a polymer. In that case the signal electrode was a microwave stripline. The relatively

Fig. 9.15 Coplanar electrode structure on x -cut (top) and z -cut (bottom) lithium niobate [51]. © 1993 IEEE



low index of refraction in the polymer (1.65) closely matched the rf index in the stripline (1.61), which yielded a 3 dB bandwidth of 40 GHz. Zhang et al. [52] have used MOCVD epitaxial growth of InGaAsP to make an electroabsorption modulator with traveling wave electrodes. The electrodes formed a coplanar waveguide (CPW) transmission line. A driving voltage of only 1.2 volts was required for an extinction ratio of 20 dB for operation at a wavelength of 1.55 μm . The modulation bandwidth was 25 GHz.

A thorough development of the theory of operation of an electro-optic waveguide modulator with traveling-wave electrodes has been done by Jaeger and Lee [53]. They demonstrated that electrodes having a microwave effective refractive index of 3.5, to match that of GaAs/InP-based semiconductors could be made with 50 and 75 Ω characteristic impedances.

Electro-optic modulators have been so widely used in various applications that they have become standard “off-the-shelf” products for many companies. Global-Spec®, The Engineering Search Engine (GlobalSpec.com) lists 47 different products from 22 suppliers for electro-optic modulators.

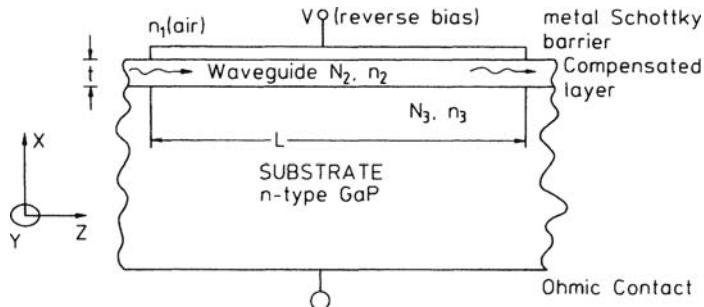
In this chapter, a number of different types of electro-optic modulators have been discussed. One of these, the Bragg diffraction type, uses an electro-optically induced grating to change the path of the optical beam. In the next chapter, we will see how surface acoustic waves can be used for the same purpose.

Problems

9.1 We wish to design a GaP electro-optic phase modulator as shown below for operation at 6300 Å wavelength.

- (a) What is the minimum thickness (t) required in the waveguiding layer if the carrier concentrations are $N_2 = 1 \times 10^{15} \text{ cm}^{-3}$ and $N_3 = 3 \times 10^{18} \text{ cm}^{-3}$?
- (b) How large a voltage (V) can be applied without producing electrical breakdown?
- (c) If this voltage is applied, how long (L) must the device be to produce a phase shift of π radians in the transmitted light wave? Assume the incident light is polarized in the Y direction. For GaP:

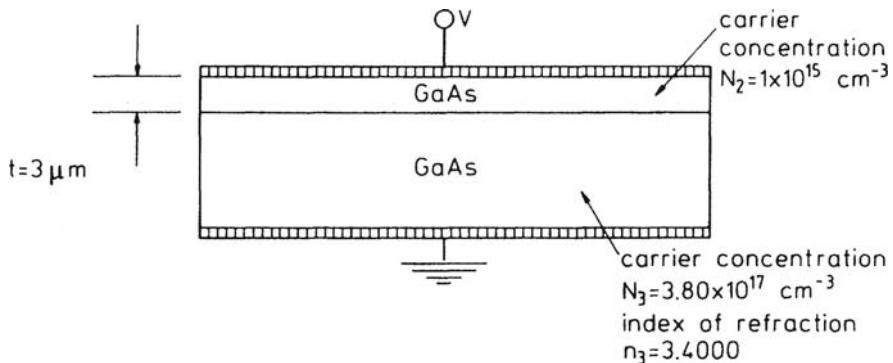
E_c (critical electric field for breakdown): 5×10^5 V/cm,
 r_{41} (electro-optic coefficient for the above orientation): 5×10^{-11} cm/V,
 m^* (effective mass): $0.013m_0$.
 $n = 3.2$.



- 9.2 In the electro-optic waveguide switch shown in the figure, how large a voltage (V) must be applied to the electrode to turn the waveguide on – i.e., to increase the index of refraction in the waveguide sufficiently so as to bring it above cutoff for the lowest order waveguide mode?

Assume:

- (1) Wavelength of light in air $\lambda_0 = 1.0 \mu\text{m}$.
- (2) Effective mass of electron $m^* = 0.08m_e$.
- (3) All of the voltage is dropped over the waveguide thickness rather than in the substrate.



- 4) Crystal orientation is such that r_{41} is the appropriate electro-optic tensor element to be used, and positive voltage (V) produces an increase in index of refraction.

- 9.3 (a) A Schottky barrier type electro-optic modulator can be used to produce which of the following types of light modulation? – phase, polarization, frequency, intensity, pulse code?
- (b) Describe the Franz-Keldysh effect.
- 9.4 In an electro-optic modulator, does the polarization of the light wave generally have an important effect on the modulation depth? Why?
- 9.5 What are the two basic conditions that must be met if optical modes in two different waveguides are to be coupled by synchronous coupling or “optical tunneling”?
- 9.6 A dual-channel GaAs waveguide modulator of the type shown in Fig. 9.5, with waveguides of $5 \mu\text{m} \times 5 \mu\text{m}$ cross-section, is used to modulate light of wavelength (in air) $\lambda_0 = 1.06 \mu\text{m}$. It has a coupling coefficient $\kappa = 1 \text{ cm}^{-1}$ for the coupling of optical energy from one channel to the other when the voltage applied to the field plates is zero. (Assume absorption loss is negligible).
- What length L must the device have to completely couple the light from channel 0 to channel 1? (Assume lossless propagation of the lowest order mode only).
 - What is the minimum voltage V_1 which must be applied to cause the light output from channel 1 to go to zero if $V_0 = 0$ and if the effective electro-optic coefficient for the geometry shown is $r = 1 \times 10^{-12} \text{ m/V}$. Assume that the substrate conductivity is so large that all of the electric field is confined to the waveguide channel.
- 9.7 In the waveguide modulator structure of Problem 9.1, with all parameters being as given in that problem:
- What is the index of refraction difference at the interface between layers 2 and 3 due to the difference in carrier concentrations?
 - What applied voltage would be required to produce a Δn of the same magnitude due to the electro-optic effect? (Assume the minimum waveguide thickness as calculated in Problem 9.1a).
 - For the special case of $N_2 = N_3 = 1 \times 10^{15} \text{ cm}^{-3}$ (i.e. a Schottky barrier formed directly on a uniformly doped substrate), would the answer to part (b) be the same?
- 9.8 (a) Sketch the shape of the electric-field distribution versus depth for the waveguide modulator structures of Problems 9.1 and 9.7c. assuming that the applied voltages have been adjusted to produce equal waveguide thickness in each case.
- (b) Would the difference in these two field distributions have a more pronounced effect on the profile of Δn versus depth in a lightly doped or more heavily doped semiconductor substrate.
- 9.9 If the maximum modulation bandwidth of an electro-optic modulator is given by

$$\Delta f = (2\pi R_L C)^{-1}$$

where R_L is the load resistance of the circuit and C is the device capacitance, shown that the power to produce a given phase change $\Delta\Phi_{EO}$ is given by

$$P = \frac{(\Delta\Phi_{EO})^2 \lambda_0^2 \varepsilon A (\Delta f) t_g}{\pi L^2 n^6 r^2}$$

where λ_0 denotes the wavelength in air, ε is the permittivity of the material, A the cross-sectional area, L the optical path length in modulator, n the index of refraction of material, r the appropriate element of the E-O tensor, and t_g the waveguide thickness.

- 9.10 To make a Franz-Keldysh type electroabsorption modulator for light of $\lambda_0 = 1.15 \mu\text{m}$ you would choose a semiconductor with what bandgap?
- 9.11 An electro-absorption modulator has an absorption coefficient $\alpha = 0.2 \text{ cm}^{-1}$ with no voltage applied, and 2 cm^{-1} with the maximum voltage of 5 volts applied. It has a length of 2 cm. This modulator is inserted into a waveguide that is initially transmitting an optical power of 500 milliwatts.
 - (a) What is its insertion loss in dB?
 - (b) What is its extinction ratio (maximum modulation depth) in dB?

Assume that the coupling losses at the input and output are negligible. Hint: Depending on how you do the problem, you may not need to use the 500 milliwatts.

References

1. D.A. Pinnow: IEEE J. **QE-6**, 223 (1970)
2. J.M. Hammer: Modulation and switching of light in dielectric waveguides, in *Integrated Optics*, T. Tamir, (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1979) p. 142
3. A. Yariv: *Quantum Electronics*, 3rd edn. (Wiley, New York 1989) pp. 298–307
4. J.F. Nye: *Physical Properties of Crystals* (Oxford University Press, New York 1957) p. 123
5. L.A. Shuvalov (ed.): *Modern Crystallography IV*, Springer Ser. Solid-State Sci., Vol. 37 (Springer, Berlin, Heidelberg 1988)
6. R.C. Alferness: Titanium-diffused lithium niobate waveguide devices, in *Guided-Wave Optoelectronics*, T. Tamir (ed.), 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) Chap. 4, in particular, pp. 155–157
7. D. Hall, A. Yariv, E. Garmire: Appl. Phys. Lett. **17**, 127 (1970)
8. I.P. Kaminov, V. Ramaswamy, R.V. Schmidt, F.H. Turner: Appl. Phys. Lett. **27**, 555 (1975)
9. I.P. Kaminov, L.W. Stultz, E.H. Turner: Appl. Phys. Lett. **27**, 555 (1975)
10. Y. Shuto, M. Amano, T. Kaino: IEEE Photon. Tech. Lett. **3**, 1003 (1991)
11. J.-M. Brosi, C. Koos, L.C. Andreani, M. Waldow, J. Leuthold, W. Freude: High-speed low-voltage electro-optic modulator with a polymer-infiltrated silicon photonic crystal waveguide, Opt. Express **16**, 4177 (2008)
12. J.C. Campbell, F.A. Blum, D.W. Shaw: Appl. Phys. Lett. **26**, 640 (1975)

13. M. Kawabe, S. Hirata, S. Namba: IEEE Trans. CAS-**26**, 1109 (1979)
14. J.I. Pamkove: *Optical Processes in Semiconductors* (Prentice Hall, Englewood Cliffs, NJ 1971) p. 29
15. B.G. Streetman: *Solid State Electronic Devices*, 4th edn. (Prentice Hall, Englewood Cliffs, NJ 1995) pp. 301–307
16. V.S. Vavilov: Sov. Phys. – Uspekhi **4**, 761 (1962)
17. F.K. Reinhart: Appl. Phys. Lett. **22**, 372 (1973)
18. Y. Node: IEEE J. LT-**4**, 1445 (1986)
19. H. Soda, K. Nakai, H. Ishikawa: High-speed and low-chirp GaInAsP/InP optical intensity modulator. *Integrated and Guided Wave Optics*, 1988 Techn. Digest Ser., Vol. 5 (Opt. Soc. Am., Washington, DC 1988) p. 28
20. A. Yariv: IEEE J. QE-**9**, 919 (1975)
21. E.A.J. Marcatili: Bell Syst. Techn. J. **48**, 2130 (1969)
22. S. Somekh, E. Garmire, A. Yariv, H.L. Garvin, R.G. Hunsperger: Appl. Phys. Lett. **22**, 46 (1973)
23. S. Somekh, E. Garmire, A. Yariv, H.L. Garvin, R.G. Hunsperger: Appl. Opt. **13**, 327 (1974)
24. H.F. Taylor: J. Appl. Phys. **44**, 3257 (1973)
25. J.C. Campbell, F.A. Blum, D.W. Shaw, K.I. Lawley: Appl. Phys. Lett. **27**, 202 (1975)
26. M. Papuchon, Y. Comberale, X. Mathieu, D.B. Ostrowsky, L. Reiber, A.M. Roy, B. Sejourne, M. Werner: Appl. Phys. Lett. **27**, 289 (1975)
27. H. Kogelnik, R.V. Schmidt: IEEE J. QE-**12**, 396 (1976)
28. Y. Zhou, W. Qiu, Y. Chen: Optica Sinica **14**, 264 (1994)
29. R.C. Alferness: Titanium diffused lithium niobate devices, in *Guided-Wave Optoelectronics*, T. Tamir (ed.), 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) Chap. 4, in particular, ps. 179, 180
30. J.J. Veselka, D.A. Herr, T.O. Murphy, L.L. Buhl, S.K. Korotky: IEEE J. LT-**7**, 908 (1989)
31. R.A. Steinberg, T.G. Giallorenzi: IEEE J. QE-**13**, 122 (1977)
32. R.A. Steinberg, T.G. Giallorenzi, R.G. Priest: Appl. Opt. **16**, 2166 (1977)
33. Y. Sikorski, R.T. Deck, A.L. Sala, B.G. Bagley: Analysis of crosstalk between single-mode rectangular optical waveguides. Opt. Eng. **39**, 2015 (2000)
34. F. Zernike: Integrated optic switch. OSA Topical Meeting on Integrated Optics, New Orleans, LA (1974)
35. W.E. Martin: Appl. Phys. Lett. **26**, 562 (1975)
36. R.A. Becker: Appl. Phys. Lett. **43**, 131 (1983)
37. C.C. Teng: Traveling-wave polymeric optical intensity modulator with more than 40 GHz of 3-dB electrical bandwidth. Appl. Phys. Lett. **60**, 1538 (1992)
38. V. Ramaswami, M.D. Divino, R.D. Standley: Appl. Phys. Lett. **32**, 644 (1978)
39. K. Izuka: *Engineering Optics*, 2nd edn., Springer Ser. Opt. Sci., Vol. 35 (Springer, Berlin, Heidelberg 1985) p. 395
40. J.M. Hammer: Modulation and switching of light in dielectric waveguides, in *Integrated Optics*, T. Tamir (ed.) 2nd edn., (Springer, Berlin, Heidelberg 1979) p. 182
41. J.M. Hammer: Appl. Phys. Lett. **18**, 147 (1971)
42. D.P. Giarrusso, J.H. Harris: Appl. Opt. **10**, 27861 (1971)
43. J.M. Hammer, D.J. Channin, M.T. Duffy: Appl. Phys. Lett. **23**, 176 (1973)
44. J.M. Hammer, W. Phillips: Appl. Phys. Lett. **24**, 545 (1974)
45. Y. Lee, S. Wang: Appl. Opt. **15**, 1565 (1976)
46. C. Xin, C.S. Tsai: Electrooptic Bragg-diffraction modulators in GaAs/AlGaAs heterostructure waveguides. IEEE J. Lightwave Technol. **6**, 809 (1988)
47. G.L. Tangonan, L. Persechini, J.F. Lotspeich, M.K. Barnoski: Appl. Opt. **17**, 3259 (1978)
48. C.S. Tsai, B. Kim, F.R. El-Akkari: IEEE J. QE-**14**, 513 (1978)
49. S.K. Sheem: Appl. Opt. **17**, 3679 (1978)
50. K. Oh, K. Park, D. Oh, H. Kim, H. Park, K. Lee: IEEE Photon. Tech. Lett. **6**, 65 (1994)

51. R.L. Jungerman, D.W. Dolfi: Lithium Niobate traveling-wave optical modulators to 50 GHz. IEEE/LEOS Topical Meeting on Optical-Microwave Interactions, Santa Barbara, CA, July 1993
52. S. Zhang, Y. Chiu, P. Abraham, J. Bowers: Traveling-wave Electroabsorption Modulator. IEEE Phot. Technol. Lett. **11**, 191 (1999)
53. N.A.F. Jaeger, Z.K.F. Lee: Slow-wave electrode for use in compound semiconductor electrooptic modulators, IEEE J. Quant. Electron. **28**, 1778 (1992)

Chapter 10

Acousto-Optic Modulators

In the preceding chapter, we have shown that modulators and switches can be made by using the electro-optic effect to produce a grating-shaped variation of the index of refraction within the waveguide. This grating structure causes diffraction of the guided optical waves, resulting in modulation or switching.

Acoustic waves can also be used to produce the desired grating pattern in the index profile. The acousto-optic effect is the change in the index of refraction caused by mechanical strain which is introduced by the passage of an acoustic-strain wave. The resulting index variation is periodic, with a wavelength equal to that of the acoustic wave. Two basic types of acousto-optic modulators are discussed in this chapter: the Bragg and the Raman-Nath configurations, which differ mainly in the interaction length between the acoustic and optical waves.

10.1 Fundamental Principles of the Acousto-Optic Effect

Mechanical strain in a solid causes a change in the index of refraction which can affect the phase of a light wave traveling in the strained medium. This photoelastic effect, as it is called, can be characterized by a fourth rank tensor (the strain-optic, or photoelastic tensor) that relates the strain tensor to the optical indicatrix, just as the electro-optic tensor characterizes changes in the indicatrix produced by an electric field. For a thorough discussion of the strain-optic tensor, see Izuka [1].

In the case of the acousto-optic effect, mechanical strain is produced in a material by the passage of an acoustic wave. That strain therefore causes a change in the index of refraction via the photoelastic effect. Pinnow [2] has shown that the change in index of refraction Δn is related to the acoustic power P_a by the expression

$$\Delta n = \sqrt{n^6 p^2 10^7 P_a / (2 Q v_a^3 A)}, \quad (10.1)$$

where n is the index of refraction in the unstrained medium, p is the appropriate element of the photoelastic tensor, P_a is the total acoustic power in Watts, ϱ is the mass density, v_a is the acoustic velocity, and A is the cross-sectional area through which the wave travels. All quantities in (10.1) are given in cgs units except for

P_a . In terms of the commonly used acousto-optic figure of merit M_2 , (10.1) can be written as

$$\Delta n = \sqrt{M_2 10^7 P_a / (2A)}, \quad (10.2)$$

where M_2 is defined by

$$M_2 \equiv n^6 p^2 / \varrho v_a^3. \quad (10.3)$$

In crystalline solids, such as are used most often for substrates in OIC's, the acousto-optic effect depends strongly on the orientation, i.e. on p . However, this effect is relatively small even for optimum choices of material and orientation. For example at a wavelength of 6328 Å, for fused quartz M_2 equals 1.51×10^{-18} s³/gm, and for LiNbO₃ it is 6.9×10^{-18} s³/gm [2]. Thus, it can be determined from (10.2) that Δn is on the order of 10^{-4} in these materials, even for an acoustic power density of 100 W/cm². Despite the small value of Δn that can be produced by an acoustic wave, the overall effect on an optical beam can be quite significant, because each small Δn produced by a strain peak of the acoustic wave can result in optical interactions that can be accumulated constructively (or destructively) if proper phase matching is provided. Thus, substantial diffraction effects can be produced.

The acousto-optic modulators and switches used in optical integrated circuits generally employ traveling acoustic waves. Thus, the grating structure that is produced in the optical index profile is actually in motion with respect to the optical beam. However, this motion has an insignificant effect on the operation of most devices. The average effect of the moving grating structure is identical to that of a stationary one, except that the m th order of diffracted light is shifted in frequency by an amount equal to $\pm m f_0$, where f_0 is the frequency of the acoustic wave. Since the acoustic frequency is typically many orders of magnitude less than the optical frequency, the effect is generally negligible in acousto-optic phase or intensity modulators and beam deflectors. However, the effect has been used to produce optical frequency division multiplexing of signals [3].

Optical wave diffraction can be produced by interaction with either bulk acoustic waves traveling in the volume of the medium, or with surface acoustic waves (SAW), traveling within roughly an acoustic wavelength of the surface [4]. Since optical waveguides are typically only a few micrometers thick, SAW modulators and switches are compatible with most OIC applications.

Regardless of whether bulk or surface acoustic waves are used, two fundamentally different types of modulation are possible. In the Raman-Nath type of modulator the optical beam is incident transversely to the acoustic beam, and the interaction length of the optical path (i.e., the width of the acoustic beam) is relatively short, so that the optical waves undergo only a simple phase grating diffraction, thus producing a set of many interference peaks in the far field pattern. If the acoustic beam is much wider, so that the optical waves undergo only a simple phase grating diffraction, thus producing a set of many interference peaks in the far field

pattern. If the acoustic beam is much wider, so that the optical waves undergo multiple rediffraction before leaving the acoustic field, a much different diffraction pattern is produced. The diffraction in this case is similar to the volume diffraction of x-rays by multiple atomic planes in a crystal, which was first observed by Bragg. In Bragg-type acousto-optic modulators, the optical beam is incident at a specific angle (the Bragg angle) to the *bars* of the acoustically-produced index grating structure, and only one diffraction lobe is observed in the far-field pattern. A more detailed description of both Raman-Nath and Bragg-type modulators and switches is given in Sections 10.2 and 10.3.

10.2 Raman-Nath-Type Modulators

The basic structure of a Raman-Nath modulator is shown in Fig. 10.1. Light passing through this device in the z -direction undergoes a phase shift given by [5]

$$\Delta\varphi = \frac{\Delta n 2\pi l}{\lambda_0} \sin\left(\frac{2\pi y}{\Lambda}\right). \quad (10.4)$$

where Δn is the acoustically produced index change, l is the interaction length, and Λ is the acoustic wavelength. The zero of the y axis is taken at the center of the incident beam. Combining (10.4) with (10.2) yields the expression

$$\Delta\varphi = \frac{2\pi}{\lambda_0} \sqrt{M_2 10^7 P_a l / 2a} \sin\left(\frac{2\pi y}{\Lambda}\right), \quad (10.5)$$

where the fact that the area A equals l multiplied by the thickness of the acoustic beam a has been used.

For Raman-Nath type diffraction to take place, the interaction length must be so short that no multiple diffraction can occur. Such is the case when the condition

$$l \ll \frac{\Lambda^2}{\lambda} \quad (10.6)$$

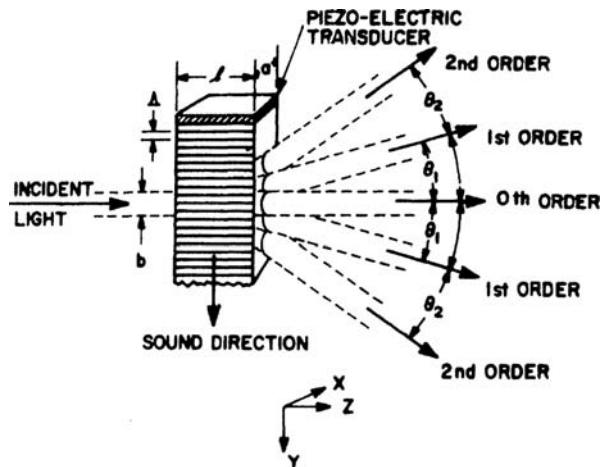
is satisfied, where λ is the optical wavelength within the modulator material. Incident light is then diffracted into a set of different orders at angles given by

$$\sin\theta = \frac{m\lambda_0}{\Lambda}, \quad m = 0, \pm 1, \pm 2, \dots \quad (10.7)$$

The intensity of these orders is given by the relation [5]

$$I/I_0 = \begin{cases} [J_m(\Delta\varphi')]^2/2, & |m| > 0 \\ [J_0(\Delta\varphi')]^2, & m = 0 \end{cases} \quad (10.8)$$

Fig. 10.1 Basic structure of a Raman-Nath type acoustooptic modulator [5]



where the J 's are the ordinary Bessel functions, I_0 is the intensity of the optical beam that is transmitted in the absence of an acoustic wave, and $\Delta\varphi'$ is the maximum value of the $\Delta\varphi$ given by (10.4), i.e.,

$$\Delta\varphi' = \frac{2\pi l \Delta n}{\lambda_0} = \frac{2\pi}{\lambda_0} \sqrt{M_2 10^7 P_a l / 2a}. \quad (10.9)$$

The output channel of a Raman-Nath modulator is usually taken to be the zeroth-order mode. In that case, the modulation index equals the fraction of the light diffracted out of the zeroth order, and is given by

$$\eta_{RN} = \frac{[I_0 - I(m=0)]}{I_0} = 1 - [J_0(\Delta\varphi')]^2. \quad (10.10)$$

Raman-Nath modulators generally have a smaller modulation index than that of comparable Bragg modulators. Also, the Raman-Nath modulator cannot be conveniently used as an optical switch, because the diffracted light is distributed over many orders, at different angles. Because of these disadvantages, the Raman-Nath modulator, while it is interesting from a theoretical standpoint, has not been often used in OIC applications. By contrast, the Bragg modulator has been widely used, as an intensity modulator, a beam deflector, and an optical switch.

10.3 Bragg-Type Modulators

For Bragg-Type diffraction to take place, the interaction length between the optical and acoustic beams must be relatively long, so that multiple diffraction can occur. A quantitative relation that expresses this condition is given by

$$l \gg \frac{\Lambda^2}{\lambda}. \quad (10.11)$$

By comparing (10.11) and (10.6) one can see that there is a transitional range of $l\lambda$ for which a composite of both Bragg and Raman-Nath diffraction occurs. However, it is usually desirable to design a modulator so that it will clearly operate in either the Bragg or Raman-Nath regime, thus enabling the input and output angles of the optical beam to be chosen for maximum efficiency.

In the case of a Bragg-type modulator, the input angle of the optical beam should optimally be the Bragg angle θ_B , given by

$$\sin \theta_B = \frac{\lambda}{2\Lambda}. \quad (10.12)$$

The diffracted (1st order) output beam emerges at an angle of $2\theta_B$ with respect to the undiffracted (0th order) beam, as shown in Fig. 10.2. Generally the output of the modulator is taken to be the zeroth order beam. In that case, the modulation depth is given by [5]

$$\frac{I_0 - I}{I_0} = \sin^2 \left(\frac{\Delta\varphi'}{2} \right), \quad (10.13)$$

where I_0 is the transmitted intensity in the absence of the acoustic beam, and I is the 0th order intensity in the presence of the acoustic beam. The maximum modulation depth, or modulation index, can be obtained by combining (10.13) and (10.9), which yields

$$\eta_B = \frac{I_0 - I}{I_0} \sin^2 \left[(\pi/\lambda_0) \sqrt{10^7 M_2 P_a l / 2a} \right]. \quad (10.14)$$

The basic modulator structures shown in Figs. 10.1 and 10.2 could be either bulk or waveguide modulators, depending on the ratio of the thickness a to the wavelength of light in the material. For a/λ very much greater than one, a bulk modulator would result. Bulk acousto-optic modulators were widely used even before the advent of integrated optics, and their use continues today. For a review of typical applications see, for example, the articles by Adler [6], Dixon [7], and Wade [8].

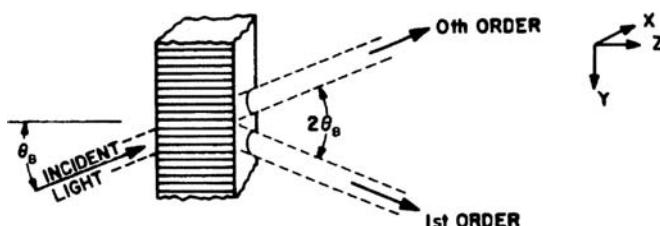


Fig. 10.2 Basic structure of a Bragg-type acousto-optic modulator [5]

Waveguide acousto-optic modulators, such as are used in OIC's, function in essentially the same manner as their bulk counterparts. However, they have the advantage of reduced drive power requirements because the optical and acoustic waves can both be confined to the same small volume.

The predicted performance of either bulk or waveguide modulators can be calculated by using (10.8) and (10.13). However, the calculation of the phase shift $\Delta\phi$ is more complicated for the case of waveguide modulators, because the optical and acoustic fields are generally not uniform over the active volume. For the case of non-uniform fields, (10.9) cannot be used to accurately determine the phase shift. Instead, an overlap integral must be calculated [9, 10]. However, (10.9) is adequate if only approximate results are desired.

The first waveguide acousto-optic modulator, as reported by Kuhn et al. [9], was a hybrid device, with the basic geometric configuration shown in Fig. 10.3. A y-cut α quartz substrate ($n = 1.54$) was used because of its relatively large piezoelectric coefficient, while a $0.8 \mu\text{m}$ thick glass film of larger index of refraction ($n = 1.73$) was sputter-deposited to form the waveguide. An interdigitated pattern of metal film electrodes was used as a transducer to launch an acoustic surface wave in the y direction. The strain pattern produced at the surface of the substrate by this wave was coupled into the optical waveguiding glass film by mechanical contact. Guided light ($\lambda_0 = 6328 \text{ \AA}$), introduced at the Bragg angle by means of a grating coupler, was observed to be diffracted through an angle $2\theta_B$. A modulation index of $\eta_B = 70\%$ was measured for an acoustic surface wave of frequency $f_a = 191 \text{ MHz}$ and wavelength $\Lambda = 16 \mu\text{m}$.

Kuhn et al. [9] chose a hybrid approach in order to use the strong piezoelectric effect in α quartz to launch the acoustic waves. Wille and Hamilton [11] also used this approach, obtaining a Bragg modulation index of 93% at a wavelength of 6328 \AA for a sputter-deposited Ta_2O_5 waveguide on an α quartz substrate, driven by a 175 mW acoustic wave with $f_a = 290 \text{ MHz}$. Omachi [12] obtained similar results with an As_2S_3 waveguide film on a LiNbO_3 substrate. At a wavelength of $1.15 \mu\text{m}$,

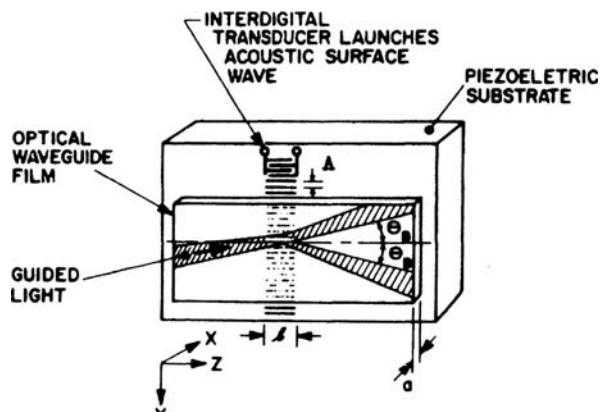


Fig. 10.3 Waveguide acousto-optic Bragg modulator structure [9]

he observed a Bragg modulation index of 93% for an acoustic power of 27 mW, at $f_a = 200$ MHz. More recently, an acoustooptic modulator has been designed using a piezoelectric film covering a multiple quantum-well structure to take advantage of the quantum-confined Stark effect [13]. (See Chap. 18 for a discussion of this phenomenon.)

While the hybrid circuit approach is obviously an effective means of obtaining high-efficiency acousto-optic modulation, it is not essential in all applications. Monolithic OIC's, in which the acoustic wave is launched directly in the waveguide material, are also possible. For example, Chubachi et al. [14] have demonstrated a Bragg acoustooptic modulator featuring a ZnO sputtered-film waveguide on a fused quartz substrate, in which the acoustic waves were launched directly into the ZnO. They obtained a modulation index of greater than 90%. Schmidt et al. [15] fabricated a Raman-Nath type modulator in an outdiffused LiNbO₃ waveguide structure. They have reported that 25 mW/MHz of acoustic power was required to completely extinguish the zeroth order transmitted beam. A review of these types of guided-wave acoustooptic Bragg modulators has been given by Tsai [16]. In that paper, the analytical and numerical techniques required for the treatment of a Bragg modulator are discussed, as well as the design parameters and procedures for wide-band Bragg modulators and deflectors.

As mentioned previously, whether Bragg or Raman-Nath type diffraction is produced depends on the interaction length shared by the optical and acoustic waves. However, the transition between these two regimes as the interaction length is increased is not abrupt. There is a "gray area" in which the diffraction follows neither the Bragg nor Raman-Nath pattern. Wang and Tarn [17] have used spatial Fourier transforms and numerical methods to theoretically analyze this "near-Bragg acoustooptic effect". Their theoretical and experimental results indicate that neither one order (as in Bragg) nor infinite orders (as in Raman-Nath) of light are significantly diffracted. Instead, four orders are observed in the near-Bragg regime, 0, +1, -1, and +2.

To conclude the discussion of basic types of acoustooptic modulators, it is appropriate to note that there are alternative methods available, in addition to the Bragg and Raman-Nath techniques. Some of these methods involve the use of acoustic waves to induce coupling between various guided modes of the waveguide, or between guided and radiation modes. Chu and Tamir [18] have developed a detailed theory of acoustic-optic interactions using the coupledmode approach, and a number of different modulators have been demonstrated that employ intermode coupling [19, 20]. Another alternative is to somehow use acoustic waves to induce a change in the optical transmission of a material, thus producing an intensity modulator. Karapinar and Gunduz [21] have done this in a homeotropically oriented nematic liquid crystal. Gryba and Lefebvre [22] have utilized the electro-absorptive effects in quantum wells and the electric field induced by surface acoustic waves to make a modulator. (See Chapter 18 for further discussion of absorption in quantum wells.) Liu et al. [23] have produced an acoustooptic Bragg-type modulator within an optical fiber by using acoustic waves to control a fiber Bragg grating. A fiber Bragg grating (FBG) is a type of distributed Bragg reflector constructed within an optical

fiber by adding a periodic variation to the refractive index of the fiber core. The FBG reflects a particular wavelength of light and transmits all others. (Distributed Bragg reflectors are discussed in Chapter 15.)

10.4 Bragg-Type Beam Deflectors and Switches

Bragg diffraction modulators are inherently capable of functioning as beam deflectors and optical switches, since the output beam exits at a different angle from the input beam. If the frequency of the acoustic wave is held constant, the optical beam can be switched through the angle $2\theta_B$ by supplying sufficient acoustic power to produce 100% diffraction from the zeroth order beam to the first-order beam. Alternatively, the frequency (wavelength) of the acoustic wave can be varied to deflect the optical beam through different angles, as given by (10.12).

In beam deflection applications, an important parameter is the number N of resolvable spots. A theoretical expression that can be used to determine N can be derived as follows. Assume that an optical beam of width b is incident on an acoustic wave over an interaction length l , as shown in Fig. 10.4. If b is much greater than Λ , so that the optical beam covers a number of acoustic wave periods, we know from the fundamental diffraction theory of a beam incident on a grating [24], that the far-field pattern contains a set of diffraction maxima having a half-power width given by

$$\Delta\theta_1 = \lambda/b, \quad (10.15)$$

with angular separation of the peaks given by

$$\Delta\theta_2 = \lambda/\Lambda. \quad (10.16)$$

When the optical beam is incident at the Bragg angle, only the first-order lobe of the diffraction pattern is significant, the other higher order have negligible intensity. The diffracted beam is thus concentrated in this peak and has maximum intensity. If the frequency (wavelength) of the acoustic wave is varied from that for which the Bragg condition is exactly satisfied, the diffracted beam is angularly scanned, but its intensity is reduced. It can be shown that the intensity follows a bell-shaped pattern with a half-power width give by [25]

$$\Delta\theta_3 = \frac{2\Lambda}{l}. \quad (10.17)$$

The number of resolvable spots is given by the ratio of the envelope width $\Delta\theta_3$ to the spot width $\Delta\theta_1$; thus, from (10.17) and (10.15)

$$N = \frac{\Delta\theta_3}{\Delta\theta_1} = \frac{2\Lambda b}{\lambda l}. \quad (10.18)$$

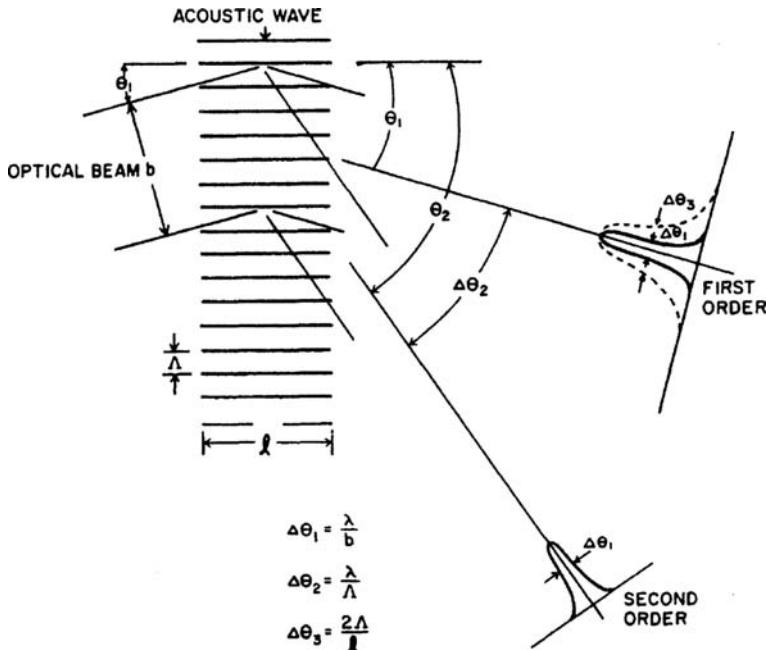


Fig. 10.4 Diffraction pattern of an optical beam incident on an acoustic wave grating. The second order (and higher) lobes would be absent in the case of $\theta_1 = \theta_B$

It must be emphasized with regard to (10.15), (10.16), (10.17) and (10.18) that angular differences $\Delta\theta_i$ ($i = 1, 2, 3$) and wavelengths Λ and λ are those measured within the medium in which the modulator is fabricated. Also, it has been assumed that the medium is optically isotropic.

It is interesting to note the equivalence between (10.17), which gives the angular deviation of the diffracted beam (from the Bragg angle) when the acoustic frequency is changed by a Δf sufficient to produce a 50% reduction in optical intensity, and (9.35), which gives the angular deviation of the input beam (from the Bragg angle) required to produce a 50% reduction in the intensity of the diffracted beam.

Another important operating characteristic of a Bragg-type beam scanner is the bandwidth Δf of the acoustic signal that it can accept. The decrease in diffracted optical beam intensity associated with angular deviation from the Bragg condition, as described above, implies an upper limit to this bandwidth. It can be shown that the maximum value of Δf corresponding to the half-power point of intensity is given by [25]

$$\Delta f_0 \cong \frac{2v_a \Lambda}{\lambda l}, \quad (10.19)$$

where v_a is the acoustic velocity. In terms of this bandwidth, the number of resolvable spots can be shown from (10.18) and (10.19) to be given by

$$N = (\Delta f_0)t, \quad (10.20)$$

where t is the transit time of the acoustic wave across the optical beam width, given by

$$t = b/v_a. \quad (10.21)$$

The limitation of bandwidth by optical factors as expressed in (10.19), (10.20) and (10.21) is not the only effect that must be considered. Very often the bandwidth is limited to significantly lower frequencies by the response time of the piezoelectric transducer that is used to launch the acoustic waves. The overall response time τ of an acoustic beam deflector is thus given by

$$\tau = \frac{1}{\Delta f_0} + \frac{1}{\Delta f_a} + t, \quad (10.22)$$

where Δf_a is the bandwidth of the acoustic transducer, and Δf_0 and t have been defined previously.

If the number of resolvable spots is desired to be much greater than one, it can be seen from (10.20) the $1/\Delta f_0$ will be much less than t . Hence, in that case, (10.22) reduces to

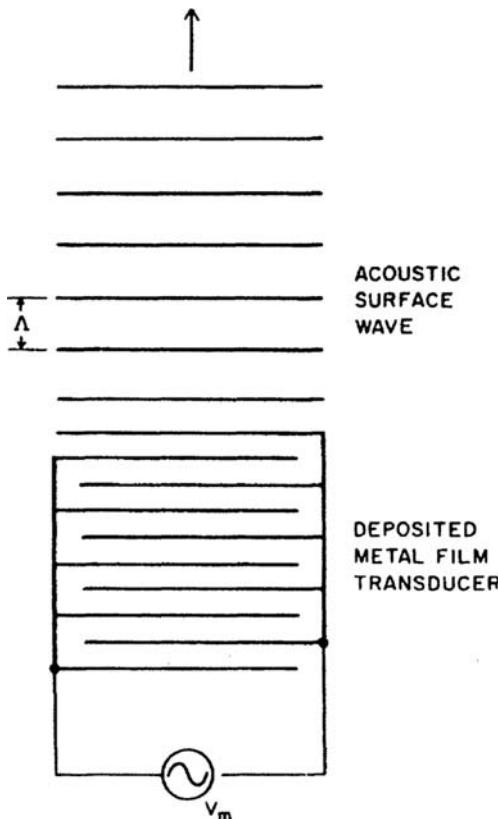
$$\tau \cong \frac{1}{\Delta f_a} + t = \frac{1}{\Delta f_a} + \frac{b}{v_a}. \quad (10.23)$$

In order to obtain maximum speed of operation, τ must be minimized by reducing t and increasing Δf_a . Note from (10.23) and (10.20), however, that there is an unavoidable trade-off between the number of resolvable spots and speed, since the transit time can be reduced by making b smaller, but that results in fewer resolvable spots. Of course, the bandwidth of the acoustic transducer Δf_a may be the limiting factor, in any case. A number of different transducer geometries designed for small τ are described in the next section.

10.5 Performance Characteristics of Acoustic-Optic Modulators and Beam Deflectors

The simplest type of acoustic transducer that can be used in a waveguide modulator consists of an interdigitated, single-periodic pattern of metal fingers deposited directly onto the waveguide surface, as shown in Fig. 10.5. Such a transducer launches a surface acoustic wave in the direction normal to its fingers. Patterns suitable for frequencies up to 1 GHz can be fabricated using standard photolithographic techniques described in Chapter 4, while interdigitated structure for higher frequencies may require the use of electron beam lithography [26, 27]. This relatively simple transducer, however, has a relatively small bandwidth compared to

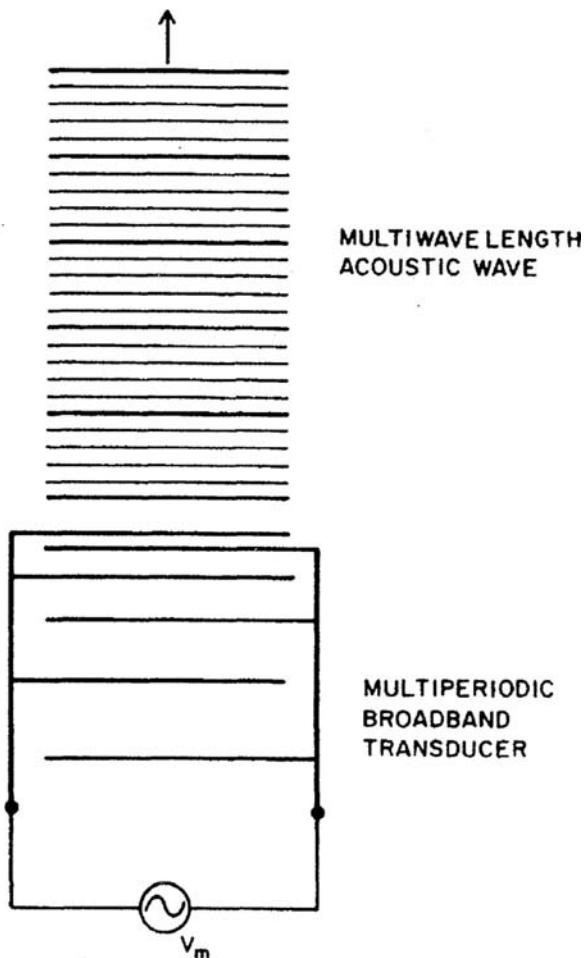
Fig. 10.5 Single-periodic interdigitated transducer.
Application of a modulation voltage V_m induces a surface acoustic wave through the piezoelectric effect



that obtainable with more sophisticated structures. For example, Tsai [28] has studied the case of a modulator for 6328 \AA light, composed of a $2 \mu\text{m}$ thick in-diffused waveguide formed in a Y-cut LiNbO_3 substrate. A transducer such as that shown in Fig. 10.5 was used, which has a center frequency of 700 MHz and an interaction length (acoustic aperture) $l = 3 \text{ mm}$. The 3 dB modulator bandwidth in this case is only 34 MHz, limited mostly by the 3 mm acoustic aperture. A larger bandwidth can be produced by reducing l . For example, a modulator such as the above, but with a waveguide thickness of $1 \mu\text{m}$, a SAW center frequency of 1 GHz, and an acoustic aperture of 0.2 mm has a bandwidth of 380 MHz [28]. However, this increased bandwidth resulting from a smaller values of l is gained at the expense of significantly reduced diffraction efficiency, see (10.14). Reduced diffraction efficiency, of course, implies that a larger acoustic drive power is required for operation.

If both wide bandwidth and high diffraction efficiency are required, more sophisticated transducer structures must be employed. One such structure is the aperiodic, or *chirp* transducer, shown in Fig. 10.6 [29, 30], in which the interdigitated spacing is gradually changed along the length of the transducer in the direction of

Fig. 10.6 Multi-periodic “chirp” interdigitated transducer



propagation of the acoustic wave. Since the transducer is most efficient in generating acoustic waves when the spacing is one-half wavelength [28], different acoustic wavelengths are optimally generated at different positions along the transducer comb, thus increasing the overall bandwidth.

Rather than using a single chirp-transducer to obtain wide bandwidth, one can use a multiple array of tilted transducers, as shown in Fig. 10.7 [31]. The individual transducers are staggered in periodicity (center frequency), and are tilted with respect to the optical beam. The tilt angle between each adjacent pair of transducers just equals the difference in the Bragg angles that correspond to their center frequencies. Thus, the composite of surface acoustic waves launched by such a transducer satisfies the Bragg condition in multiple frequency ranges, and thereby yields a much wider bandwidth than that which could be obtained with a single transducer.

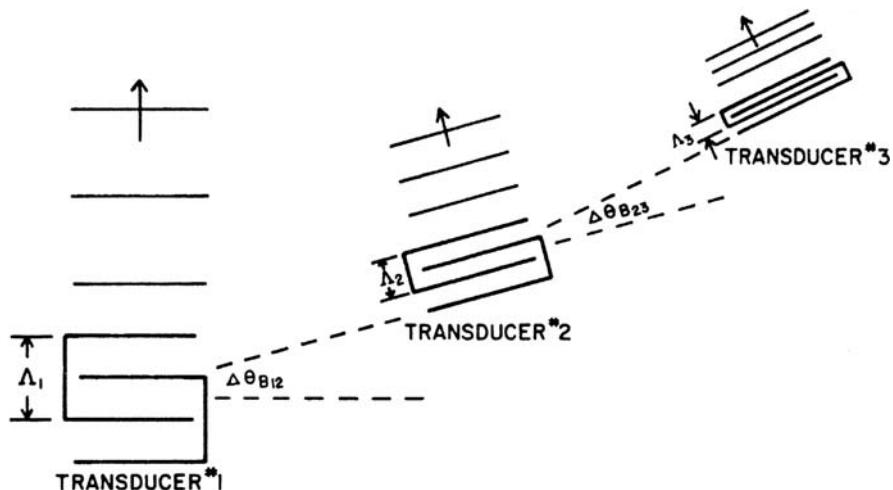


Fig. 10.7 Multi-periodic transducers in a tilted array for increased bandwidth

The electrical drive power must be coupled to each transducer through a suitable matching network, but the transducers are then driven in parallel through a power divider.

Lee et al. [32] have fabricated a modulator with multiple tilted transducers in a Ti-diffused waveguide in Y-cut LiNbO₃. The transducers had center frequencies of 380, 520, 703 and 950 MHz. The TE₀ mode from a He-Ne laser at 6328 Å was modulated over a 680 MHz bandwidth with a diffraction efficiency of 8% at a total rf drive power of 0.8 W. The overall drive power of 1.17 mW/MHz was reasonably low, even though the conversion efficiency of the 950 MHz transducer was very poor (-15 dB). Tsai [28] predicted that, when more efficient high frequency transducers have been fabricated, modulators of this type should be capable of a 1 GHz bandwidth with 50% diffraction efficiency at an electrical drive power of 1 mW/MHz.

Multiple tilted transducers are very effective because they permit the Bragg condition to be satisfied over a large acoustic aperture and over a wide frequency range. Another approach for achieving the same objective is to use a phased-array transducer set, as shown in Fig. 10.8 [33]. In this case, the individual transducer elements have identical center frequencies and parallel propagation axes. However, they are arranged in a stepped configuration that results in a variable phase shift between adjacent surface acoustic waves as the SAW frequency is varied. Because of this phase shift and resultant wave interference, a scanning acoustic wavefront is generated, in the same manner that a scanning radar beam is generated by a phased-array antenna. Scanning of the wavefront produces a wide aperture acoustic beam that tracks the Bragg condition over a large frequency range. Nguyen and Tsai [33] have fabricated a modulator of this type in a 7 μm thick outdiffused LiNbO₃ waveguide. A six element array with a center frequency of 325 MHz and a total acoustic aperture of 10.44 mm required only 3.5 mW of acoustic drive power (68 mW of

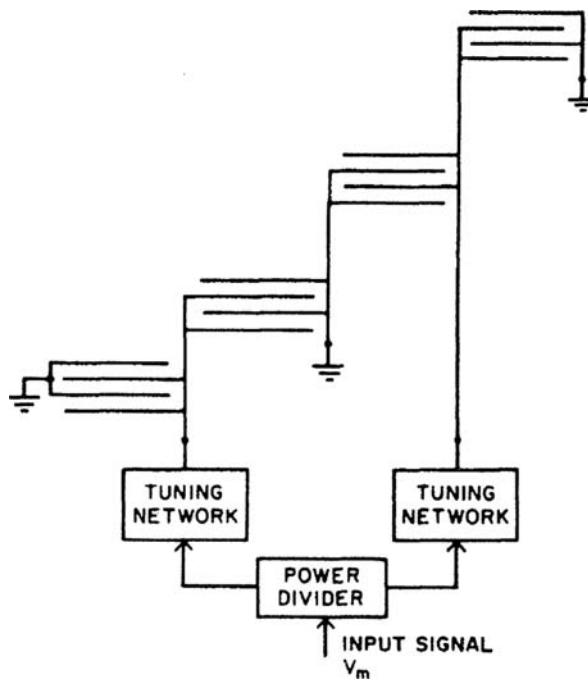


Fig. 10.8 Phased array transducers [33]

electrical power) to diffract 50% of the light, over a bandwidth of 112 MHz. As in the case of other multi-element wide bandwidth acousto-optic modulators that have been described, Δf was limited by the transducer bandwidth rather than by any optical phenomena.

Acousto-optic modulators and beam deflectors are among the more advanced integrated optic devices. They have not only been developed to a high degree of sophistication in the laboratory, but have also been employed in a number of practical applications, such as the rf spectrum analyzer [34], which is described in Chapter 20, and the DOC II 32-bit digital optical computer [35], which employs multichannel acousto-optic modulators.

10.6 Acousto-Optic Frequency Shifters

In addition to their use in modulators and beam deflectors, acousto-optic devices can also be employed to shift the frequency of an optical wave. As mentioned in Section 10.1, the frequency shifting effect of an acousto-optic modulator is so small that it is usually neglected when considering modulators and beam shifters. However, even this slight shift in frequency can be significant enough to be useful in certain applications.

For example, the standing wave surface acoustic wave optical modulator (SWSAWOM) which is shown in Fig. 10.9 can be used to produce frequency shifts of an optical carrier wave that are sufficiently large to permit frequency division multiplexing of a number of information signal channels onto an optical beam [3]. In this device the light from a semiconductor laser is butt coupled into a Ti diffused waveguide formed in a LiNbO_3 substrate and is then collimated by a geodesic lens before passing through a SWSAWOM. The SWSAWOM features two surface wave acousto-optic modulator (SWAOM) transducers which launch surface waves in opposite directions. The acoustic waves traveling in one direction cause an upward shift in the frequency of the optical waves while the acoustic waves traveling in the opposite direction produce a downward shift in the frequency of the optical waves. Thus the emerging optical beam has frequency components at $f_0 + f_m$ and also $f_0 - f_m$, where f_m is the frequency at which the transducers are driven. Thus, when the optical beam is refocused by another geodesic lens and directed to a square law detector (possibly over an interconnecting optical fiber), the two components of the optical beam beat together to produce a beam with a modulated subcarrier frequency of $f_c = 2f_m$. A number of such modulators can be combined in an optical integrated circuit as shown in Fig. 10.10 to produce a frequency division multiplexed optical beam with a number of subcarrier channels at frequencies f_{c1} , f_{c2} , f_{c3} , etc. Each one of these channels separately carries the particular information signal that is impressed on its diode laser by modulating the laser drive current. These signals can be sorted out at the receiving end by conventional microwave

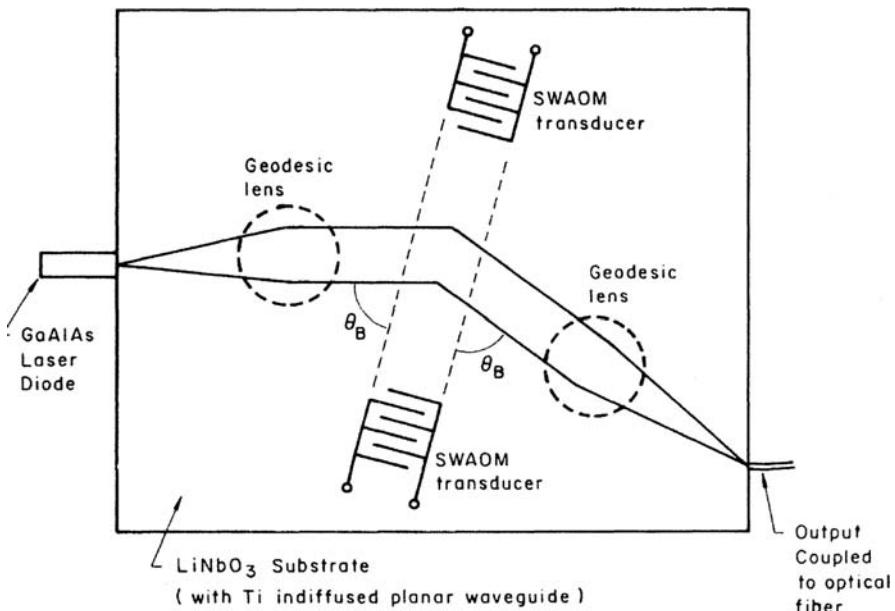


Fig. 10.9 A SWSAWOM for optical frequency shifting

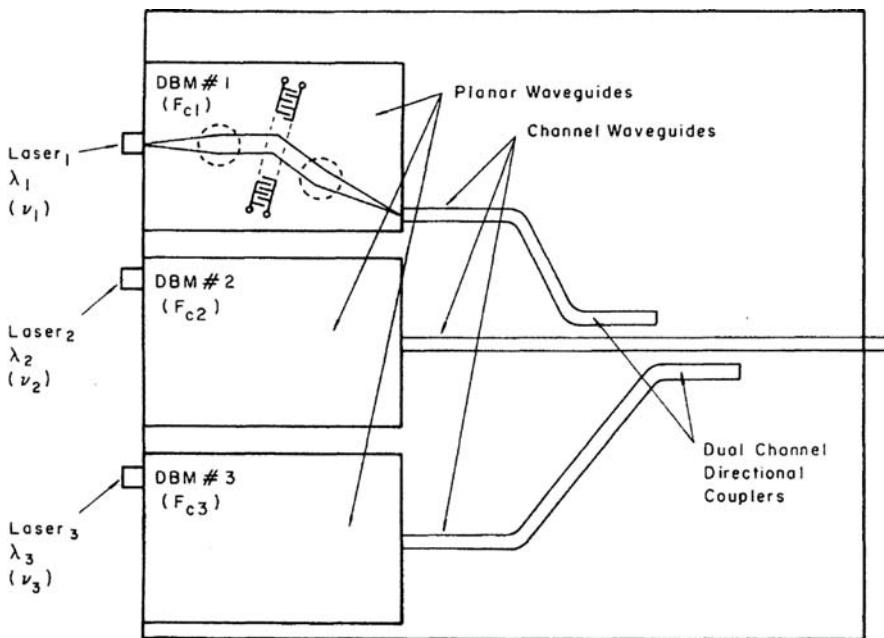


Fig. 10.10 An optical frequency division multiplexer

filtering methods. A SWSAWOM of the type shown in Fig. 10.9 has been fabricated to operate at a subcarrier frequency of $f_c = 600$ MHz [36]. In this case the transducers were driven at a frequency of 300 MHz.

Another application that uses the frequency (phase) shifting capability of the acousto-optic modulator is the Mach-Zehnder interferometer described by Bonnotte et al. [37]. The interferometer is used to measure the phase difference between the lightwaves in its two arms. In that device, which is shown in Fig. 10.11, a surface acoustic wave passes through the reference arm of a Mach-Zehnder interferometer, causing a phase shift modulation in the lightwaves in that arm relative to those in the measurement arm. An acoustic absorber prevents the acoustic waves from reaching the measurement arm.

A ZnO thin-film transducer is integrated with the silicon substrate to improve the piezoelectric coupling. The transducer is driven by a sinusoidal, 48 MHz modulation signal, the purpose of which is to facilitate measurement of the phase change $\Delta\varphi$. By comparing the amplitudes of the modulation signal of frequency f_0 and its harmonics $2f_0$ and $3f_0$ as observed on a spectrum analyzer, one can accurately determine $\Delta\varphi$.

The waveguides that were used to make the interferometer consisted of an SiON planar guide capped with a SiO_2 ridge as shown in Fig. 10.12

This strip-loaded waveguide operated at a wavelength of 633 nm and guided a single TE_{00} mode for guide widths up to 4.5–5 μm .

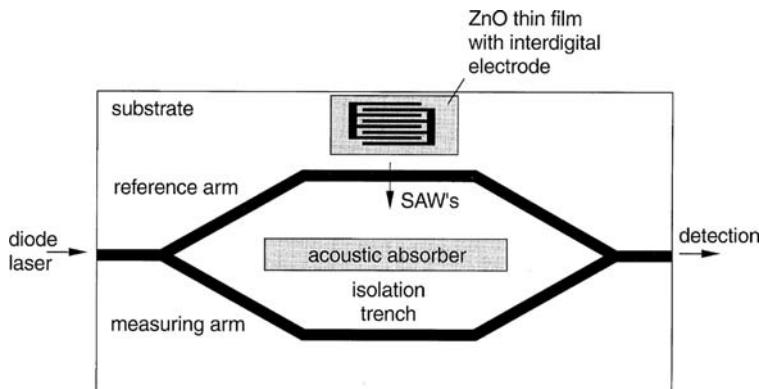


Fig. 10.11 Integrated Mach-Zehnder interferometer with phase modulator [37] © 1999 IEEE

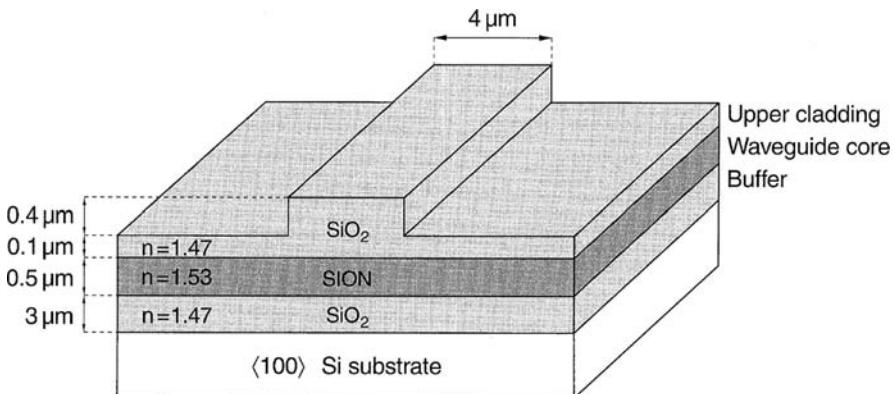


Fig. 10.12 Cross-section of the waveguide structure [37] © 1999 IEEE

Like their electro-optic counterparts, acousto-optic modulators also have become “off-the-shelf” products, with a number of different suppliers in the marketplace.

Problems

- 10.1 Show that, in general, a properly designed waveguide light modulator requires less electrical driving power than a bulk modulator of the same type (Bragg, Raman-Nath, etc.) in the same material.
- 10.2 What change in index of refraction can be induced in a quartz waveguide 3 μm thick and 100 μm wide by an acoustic power of 1 W?
- 10.3 A Bragg modulator is formed in a LiNbO_3 planar waveguide which is capable of propagating only the lowest order mode for light of 6328 Å (vacuum) wavelength. This mode has $\beta = 2.085 \times 10^5 \text{ cm}^{-1}$. The bulk index of refrac-

tion for LiNbO_3 at this wavelength is 2.295. The acoustic wavelength in the waveguide is $2.5 \mu\text{m}$, the optical beam width 4.0 mm , and the interaction length 2.0 mm . Obtain the following:

- (a) The angle (in degrees, with respect to the acoustic wave propagation direction) at which the optical beam must be introduced so as to obtain maximum diffraction efficiency.
 - (b) If the input optical beam is a uniform plane wave, what is the angular divergence (in degrees) between the half-power points of the diffracted beam, and at what angle does it leave the modulator?
 - (c) Make a sketch of the modulator, labeling the angles found in a) and b). (Do not attempt to draw the angles to scale.)
- 10.4 In a given hybrid OIC, a GaAs waveguide ($n = 3.59$) is butt-coupled to a LiNbO_3 Bragg modulator ($n_{\text{Bulk}} = 2.2$). (i) If an optical beam of vacuum wavelength $\lambda_0 = 1.15 \mu\text{m}$ is traveling in the GaAs waveguide, incident on the modulator at an angle of 93° relative to the direction of propagation of the acoustic waves, for what acoustic wavelength Λ will the modulator produce the maximum output optical intensity? (ii) At what angle will the optical beam be traveling once it has left the modulator (in air)?
- 10.5 Given a Bragg modulator with $l = 50 \mu\text{m}$, in which the acoustic velocity is 250 m/s , could the modulator be used to impress a 100 MHz bandwidth signal on an optical beam from a He-Ne laser ($\lambda_0 = 0.6328 \mu\text{m}$)?
- 10.6 An acousto-optic, Bragg-type modulator/beam deflector is fabricated in a waveguide material which has refractive index $n = 2.2$ and an acoustic velocity $v_a = 2800 \text{ m/s}$.
- (a) If the modulator is to be operated at an acoustic frequency of 100 MHz what should be the spacing between the interdigitated metal fingers of the acoustic transducer? Draw a sketch of the transducer fingers and label the spacing for clarity.
 - (b) What is the angle through which the optical beam will be deflected if the light source is a He-Ne laser with wavelength $\lambda_0 = 6328 \text{ \AA}$?
- 10.7 A Bragg-type acousto-optic beam deflector is built in a crystalline material in which the acoustic velocity is 2500 m/s . The crystal has a refractive index $n = 1.6$.
- (a) What is the acoustic wavelength required to deflect a He-Ne laser beam ($\lambda_0 = 6328 \text{ \AA}$) at a Bragg angle of 1° ?
 - (b) What is the corresponding acoustic frequency?
 - (c) Sketch an interdigitated metal-finger transducer such as could be deposited on the surface of the crystal to launch a surface acoustic wave with the wavelength calculated in (a). Show how the voltage would be applied to the fingers and label the acoustic wavelength with an arrow between the appropriate fingers.

10.8 An acousto-optic, Bragg-deflection-type RF spectrum analyzer is made in a planar waveguide in LiNbO_3 with an index $n = 2.0$. Signals with modulation frequencies ranging from 50 to 500 MHz are analyzed by using them to drive the acousto-optic transducer which produces a traveling acoustic wave in the LiNbO_3 with a velocity $v_a = 3000 \text{ m/s}$. The optical beam has a wavelength in air of $\lambda_0 = 6328 \text{ \AA}$ and a width $b = 1 \text{ mm}$. The interaction length $\ell = 5 \text{ mm}$.

- (a) What is the angular deflection (in degrees) when the modulation frequency is 500 MHz?
- (b) What is the number of resolvable spots (the ratio of the envelope width to the spot width) at a frequency of 500 MHz?

10.9 Repeat Problem 10.8 for a modulation frequency = 100 MHz.

10.10 A Bragg-type modulator made in fused quartz is used to modulate light of wavelength = 6328 \AA . The transmitted intensity in the absence of the acoustic beam is 200 mW/cm^2 and the 0th order intensity in the presence of the beam is 175 mW . The length of the modulator in the direction parallel to the optical beam is 1 cm and the thickness of the modulator is 1 mm. If the power of the acoustic beam is 500 mW, what is the maximum modulation depth?

References

1. K. Izuka: *Engineering Optics*, 3rd edn., Springer Ser. Opt. Sci., Vol. 35 (Springer, Berlin, Heidelberg 2008)
2. D. Pinnow: IEEE J. **QE-6**, 223 (1970)
3. C.S. Ih, R.G. Hunsperger, J.J. Kramer, R. Tian, X. Wang, K. Kiss, J. Butler: SPIE Proc. **876**, 30 (1988)
4. K-Y. Hashimoto: *Surface Acoustic Wave Devices in Telecommunications: Modelling and Simulation* (Springer, Berlin, Heidelberg, 2000)
5. J.M. Hammer: Modulation and switching of light in dielectric waveguides, in *Integrated Optics*, T. Tamir (ed.), 2nd edn., Topics Appl. Phys., Vol. 7 (Springer, Berlin, Heidelberg 1982) p. 155
6. R. Adler: IEEE Spectrum **4**, 42 (May 1967)
7. R.W. Dixon: J. Appl. Phys. **38**, 5149 (1967)
8. G. Wade: Bulk-wave acousto-optic Bragg diffraction, in *Guided-wave Acousto-Optic Interactions, Devices and Applications*, C. Tsai (ed.), Springer Ser. Electron. Photon., Vol. 23 (Springer, Berlin, Heidelberg 1990) Chap. 2
9. L. Kuhn, M.L. Dakss, P.F. Heidrich, B.A. Scott: Appl. Phys. Lett. **17**, 265 (1970)
10. T.G. Gialorenzi, A.F. Milton: J. Appl. Phys. **45**, 1762 (1974)
11. D.A. Wille, M.C. Hamilton: Appl. Phys. Lett. **24**, 159 (1974)
12. Y. Omachi: J. Appl. Phys. **44**, 3928 (1973)
13. T. Gryba, J. Lefebvre, J. Gazalen: *Proc. Ultrasonics Int'l Conf. 1993* (Butterworth-Heinemann, London 1993) pp. 73–76
14. N. Chubachi, J. Kushibiki, Y. Kikuchi: Electron. Lett. **9**, 193 (1973)
15. R.V. Schmidt, I.P. Kaminow, J.R. Carruthers: Appl. Phys. Lett. **23**, 417 (1973)
16. C. Tsai: Guided-wave acousto-optic Bragg modulators for wide-band integrated optic communications and signal processing, IEEE Trans. Circ. Syst. **26**, 1072 (1979)

17. C.C. Wang, C.W. Tarn: Theoretical and experimental analysis of the near-Bragg acousto-optic effect, *Opt. Eng.* **37**, 208 (1998)
18. R.S. Chu, T. Tamir: *IEEE Trans. MTT* **18**, 486 (1970)
19. L. Kuhn, P. Heidrich, E. Lean: *Appl. Phys. Lett.* **19**, 428 (1971)
20. M.L. Shah: *Appl. Phys. Lett.* **23**, 75 (1973)
21. R. Karapinar, E. Gunduz: *Opt. Commun.* **105**, 29 (1994)
22. T. Gryba, J. Lefebvre: *IEE Proc. Optoelectronics* **141**, 62 (1994)
23. W.F. Liu, P.S.J. Russell, L.Dong: Acousto-optic superlattice modulator using a fiber Bragg grating, *Opt. Lett.* **22**, 1515 (1997)
24. A. Nussbaum, R.A. Phillips: *Contemporary Optics for Scientists and Engineers* (Prentice Hall, Englewood Cliffs, NJ 1976) pp. 248–255
25. E.G.H. Lean: Acousto-optical interactions, in *Introduction to Integrated Optics*, M.K. Barnoski (ed.) (Plenum, New York 1974) p. 458
26. S.A. Campbell: *Fabrication Engineering at the Micro- and Nanoscale*, 3rd edn. (Oxford, New York, 2008) pp. 227–234
27. S.M. Sze: *VLSI Technology*, 2nd edn. (McGraw-Hill, New York 1988) pp. 450–457
28. C.S. Tsai: *IEEE Trans. CAS* **26**, 1072 (1979)
29. R.H. Tancrell, M.G. Holland: *IEEE Proc.* **59**, 393 (1971)
30. W.R. Smith, H.M. Gerard, W.R. Jones: *IEEE Trans. MTT* **20**, 458 (1972)
31. C.S. Tsai, M.A. Alhaider, Le T. Nguyen, B. Kim: *IEEE Proc.* **64**, 318 (1976)
32. C.C. Lee, B. Kim, C.S. Tsai: *IEEE J. QE* **15**, 1166 (1979)
33. L.T. Nguyen, C.S. Tsai: *Appl. Opt.* **16**, 1297 (1977)
34. D. Mergerian, E.G. Malarkey: *Microwave J.* **23**, 37 (1980)
35. R. Stone, F. Zoise, P. Guifoyle: *SPIE Proc.* **1563**, 267 (1990)
36. K. Kiss, X. Wang, C.S. Ih, R.G. Hunsperger: Novel integrated-optic modulator for optical communications. OSA Opticon'88 (Santa Clara, CA 1988) Paper MK3
37. E. Bonnotte, C. Gorecki, H. Toshiyoshi, H. Kawakatsu, H. Fujita, K. Worhoff, K. Hashimoto: Guided-Wave Acoustooptic Interaction with Phase Modulation in a ZnO Thin-Film Transducer on a Si-Based Integrated Mach-Zehnder Interferometer. *IEEE J. Lightwave Tech.* **17**, 35 (1999)

Chapter 11

Basic Principles of Light Emission in Semiconductors

This chapter is the first in a sequence of five chapters that describe the light sources used most often in integrated-optic applications. Because of their convenience, gas lasers are frequently used in the laboratory to evaluate waveguides or other integrated-optic devices; however, semiconductor lasers and light-emitting diodes are the only practical light sources for use in optical integrated circuits, due to their small size and compatibility with monolithic (or hybrid) integration. Also, LED's and laser diodes are widely used in fiberoptic applications, because they can be modulated at high frequencies and can be efficiently coupled to the micrometer-size wave-guiding core of the fiber.

Since its invention in the early nineteen sixties, the semiconductor laser has matured into a highly sophisticated, efficient light source. In order to understand the subtleties of its operation, and how to best use it in a given application, one must be familiar with some of the basic principles of light generation (and absorption) in solids. However, a thorough understanding of the quantum mechanical theory of the interaction of radiation and matter is not necessary. Hence, in this chapter, light generation and absorption in a semiconductor are described in terms of a simplified model that avoids much of the mathematical complexity of the full theory, but nevertheless incorporates all of the essential features of the phenomena that are involved. In this model, both electrons (or holes) and photons of light are described in terms of their energy, mass and momentum within the crystalline material, taking into account the effect of the lattice atoms on these quantities. When this is done, optical processes such as emission and absorption can be conveniently seen to be regulated by the familiar laws of conservation of energy and momentum.

11.1 A Microscopic Model for Light Generation and Absorption in a Crystalline Solid

11.1.1 Basic Definitions

In considering the interaction of light with the electrons of a semiconductor, a convenient model to use is a microscopic (as opposed to macroscopic) *crystal momentum*

model [1] in which the electron is considered to have a crystal momentum defined by

$$\mathbf{p} = \hbar\mathbf{k}, \quad (11.1)$$

where \mathbf{k} is the wavevector of the electron state (Plank's constant h equals 6.624×10^{-27} ergs and $\hbar = h/2\pi$). The vector \mathbf{p} is not the classical momentum of a free electron mv , where m denotes the mass and v the velocity. The crystal momentum \mathbf{p} includes the effect on the electron of the atoms of the crystal, and it is in general related to the electron energy and velocity in a rather complicated way. In the equations of motion of an electron in a crystal, \mathbf{p} plays the role of momentum and the effective mass m^* plays that of mass by analogy with the classical equations of motion of a free electron in a vacuum. Crystal momentum and effective mass can be defined for holes in the valence band just as is done for electrons in the conduction band. However, in general \mathbf{p} and m^* for electrons will be different from those for holes in a given material.

Interacting with the electrons and holes in a semiconductor, there can be photons of light. These discrete quanta, or relatively localized units of optical energy, maintain their identity as units throughout the processes of emission, transmission, reflection, diffraction, absorption, etc. The photon is described in a microscopic model by defining its mass and momentum as follows. First, consider a photon of light in a vacuum. The energy E of the photon is given by

$$E = h\nu, \quad (11.2)$$

where ν is the frequency of the radiation in s^{-1} . If (11.2) is combined with the well known relation

$$c = \nu\lambda_0, \quad (11.3)$$

the result is that the energy of the photon is given by

$$E = \frac{hc}{\lambda_0}, \quad (11.4)$$

where c is the velocity and λ_0 is the wavelength of light in a vacuum. It is often convenient to remember that (11.4) becomes

$$E = \frac{1.24}{\lambda_0} \quad (11.5)$$

in the commonly used unit system in which E is given in electron volts (eV) and wavelength is in micrometers.

From the basic equation giving the energy of a photon (11.4), it is possible to designate its *mass* and *momentum*. It is known from relativistic theory that, in general, the energy of a particle and its rest mass are related by [2]

$$E = mc^2. \quad (11.6)$$

Combining (11.4) and (11.6) yields

$$E = mc^2 = \frac{hc}{\lambda_0}. \quad (11.7)$$

Although a photon has no rest mass, one can identify

$$m = \frac{h}{c\lambda_0}. \quad (11.8)$$

By analogy with the classical momentum p , which is equal to the product of mass times velocity, the momentum of a photon is given by

$$\mathbf{p} = mc\mathbf{u}, \quad (11.9)$$

where \mathbf{u} is a unit vector in the direction of travel of the photon. Combining (11.9) and (11.8), the momentum can also be written as

$$\mathbf{p} = \frac{h}{\lambda_0}\mathbf{u} = \hbar\mathbf{k}, \quad (11.10)$$

where \mathbf{k} is the wavevector which is equal to $(2\pi/\lambda_0)\mathbf{u}$.

It can thus be seen from (11.2), (11.3), (11.4), (11.5), (11.6), (11.7), (11.8), (11.9) and (11.10) that, in this microscopic model, photons may be described equivalently by either their energy, wavelength, frequency, mass or momentum. Note that, in the case of the first four, the direction of travel must also be specified in order to completely describe the photon. The phase of photons is also important in cases involving coherent radiation, as is described in following sections of this chapter. Equations (11.2), (11.3), (11.4), (11.5), (11.6), (11.7), (11.8), (11.9) and (11.10) were developed for the case of a photons in a vacuum. However, to use them to describe photons in a solid it is merely necessary to replace c and λ_0 respectively with v and λ , the velocity and wavelength of light in the material, respectively.

The interaction of photons of radiation with electrons and holes in a semiconductor can be considered as *particle* interactions by using the definitions presented in the preceding paragraphs. In that case, the usual rules of conservation of energy and momentum regulate the optical processes.

11.1.2 Conservation of Energy and Momentum

The effect of energy and momentum conservation can be conveniently illustrated with regard to the phenomenon of optical absorption. In Chapter 6, absorption loss was considered from the macroscopic point of view, being characterized by an absorption coefficient α . The microscopic model described in Section 11.1.1 can be used to explain further details of the absorption process. The strongest absorption in semiconductors occurs through electronic excitation, when an electron in a given energy level absorbs a photon and makes a transition to a higher energy state. Electronic transitions are subject to certain selection rules – the most basic being that *energy* and *momentum* of the electron and photon must be conserved. (From a theoretical viewpoint, it is a conservation of wavevector \mathbf{k} , resulting from periodicity of the crystal structure, which leads to conservation of momentum, defined as $\hbar\mathbf{k}$.) Thus, if only photons and electrons are involved,

$$E_i + h\nu_{\text{phot}} = E_f, \quad (11.11)$$

and

$$\mathbf{p}_i + \mathbf{p}_{\text{phot}} = \mathbf{p}_f \text{ or } \mathbf{k}_i + \frac{2\pi}{\lambda_{\text{phot}}} \mathbf{u} = \mathbf{k}_f, \quad (11.12)$$

where the subscripts i and f refer to the initial and final states of the electron. The momentum of a photon in the visible or infrared range of the spectrum is very much smaller than the momentum of a thermally excited electron; hence the momentum selection rule is approximately $\mathbf{k}_i = \mathbf{k}_f$, and the electron experiences practically no change of momentum. This type of a transition, involving only an electron and photon is called a *direct* transition. It appears on the E versus \mathbf{k} diagram as a *vertical* transition, as shown in Fig. 11.1. Note that, in general, the conduction band minimum and the valence band maximum do not have to occur at the same point in \mathbf{k} space. If they do, the material is said to have a direct bandgap; if they do not, it is an indirect bandgap material.

Even though a photon cannot transfer significant momentum to an electron, it is possible to have electronic transition in which \mathbf{k} changes. In these processes, a

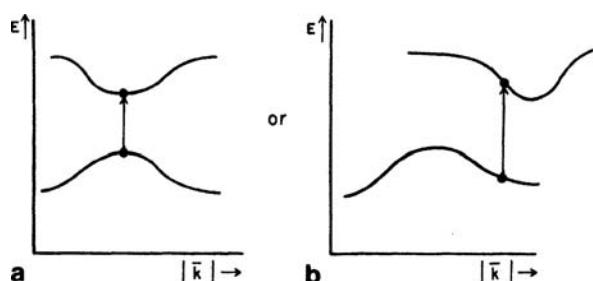


Fig. 11.1 a, b Direct absorptive electronic transitions; **a** in a direct bandgap material; **b** in an indirect bandgap material

photon also takes part, or equivalently speaking, momentum is transferred to or absorbed from the lattice atoms. The vibrational energy of the lattice atoms is quantized (as the light energy) into phonons of energy $\hbar\omega$ (ω : angular frequency in rad/s). The phonons are characterized by a wavevector \mathbf{q} . The conditions restricting an absorptive transition in which a photon, electron and phonon are involved, are then given by

$$\mathbf{k}_i \pm \mathbf{q} = \mathbf{k}_f, \quad (11.13)$$

and

$$E_i + h\nu_{\text{phot}} \pm h\omega_{\text{phon}} = E_f, \quad (11.14)$$

where the wavevector $\mathbf{k} = (2\pi/\lambda_0)\mathbf{u}$ of the photon has been neglected in (11.13). An E versus \mathbf{k} diagram plot of the process is shown in Fig. 11.2. An electron absorbs a photon and at the same time absorbs or emits a phonon. Such transitions are called *indirect*, and appear as diagonal transitions on the E versus \mathbf{k} diagram, as shown in Fig. 11.2. Since both a photon and a phonon are necessary for an indirect transition, while only a photon is required to cause a direct transition, the direct bandgaps are more optically active than those with indirect bandgaps. This feature is particularly important in the case of light emitters, as is described later in this chapter.

The diagrams of Figs. 11.1 and 11.2 shown only interband transitions, between valence and conduction bands; however, both direct and indirect transitions also can take place within a band (intraband) or between energy states introduced by dopant atoms and/or defects. In all cases, the principles of conservation of energy and momentum (wavevector) apply. These types of absorption transitions are shown in Fig. 11.3. Intraband absorption can occur by either electrons in the conduction band or by holes in the valence band, and hence is called *free carrier* absorption, as mentioned previously in Chapter 6. Free carrier absorption is usually taken to include transitions of electrons from donor states to the conduction band and transitions of holes from acceptor states to the valence band.

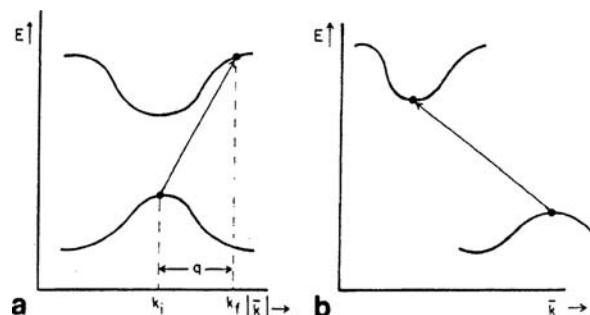
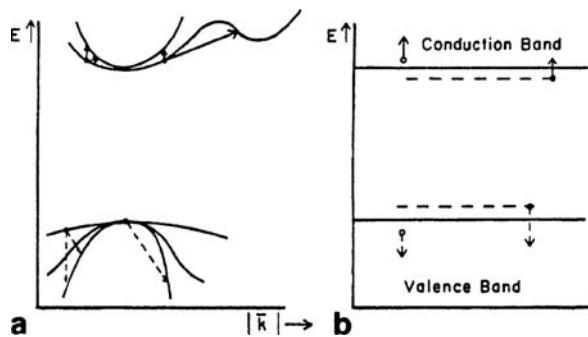


Fig. 11.2 a, b Indirect absorptive electronic transition: **a** in a direct bandgap material; **b** in an indirect bandgap material

Fig. 11.3 a, b Free carrier absorption. Since this is an electron energy diagram, increasing energy for holes corresponds to moving downward on the diagram



The effects of energy and momentum conservation that have been illustrated by reference to the optical absorption process apply equally well to the generation of photons in a semiconductor. In fact, such considerations are even more important in the case of light emission, as is discussed in more detail in Section 11.2.

11.2 Light Emission in Semiconductors

Photon generation in a semiconductor usually results from carrier recombination. The generation process can be divided into two types, spontaneous emission, in which a hole and electron randomly recombine, and stimulated emission, in which a hole and electron are stimulated to recombine by an already existing photon.

11.2.1 Spontaneous Emission

In absorption, an electron absorbs a photon and makes a transition from a higher energy state. In emission, just the opposite occurs. An electron makes a transition from a higher energy states to a lower one, and in the process loses energy that is emitted as a photon of light. The significant light emitting transitions in a semiconductor are interband transitions, in that they occur between conduction and valence bands and/or certain states within the bandgap caused by impurity dopants or defects. Since the bandgaps of semiconductors are typically a few tenths eV to a few eV, the emitted wavelengths are generally in the infrared part of the spectrum, and correspond roughly to the absorption edge wavelength of the semiconductor. As in the case of absorption, certain selection rules limit the possible emissive transitions. Conservation of energy and momentum (or wavevector) are required. Thus, for a direct transition

$$E_i - E_f = h\nu_{\text{phot}} \quad (11.15)$$

and

$$\mathbf{k}_i - \mathbf{k}_f = \frac{2\pi}{\lambda_{\text{phot}}} \mathbf{u}, \quad (11.16)$$

where \mathbf{u} is a unit vector specifying direction only. Note that for a direct transition $2\pi/\lambda_{\text{phot}} \ll |\mathbf{k}_i|$ and $|\mathbf{k}_f|$ so that $|\mathbf{k}| \approx |\mathbf{k}_f|$, as for a direct absorptive transition. Emission of a photon can also occur in an indirect transition, in which case conservation of momentum requires

$$\mathbf{k}_i - \mathbf{k}_f \pm \mathbf{q} = \frac{2\pi}{\lambda_{\text{phot}}} \mathbf{u}, \quad (11.17)$$

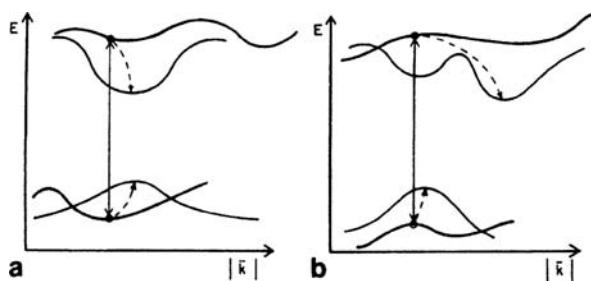
where \mathbf{q} is the wavevector of an absorbed or emitted phonon. As in the case of absorption, the necessity of having a phonon present greatly reduces the probability of a transition and hence of the generation of a photon.

A third requirement for an emissive transition is that there must be a filled upper, or initial, energy state and a corresponding empty lower, final, energy state, with a difference in energy between them equal to the energy of the photon to be emitted. This rule seems obvious and trivial, and plays little role in the case of absorption, because the absorbing material is usually near thermal equilibrium, and consequently there are always filled lower states and empty upper states at the proper levels. (Except in the case of less-than bandgap radiation, for which there are no upper states available, and hence absorption can not occur.) In the case of emission, the preceding rule is much more important. For example, for an intrinsic semiconductor at thermal equilibrium, there are relatively few holes in the valence band and electrons in the conduction band at room temperature and below. In the case of a doped semiconductor, there may be many electrons in the conduction band, as for n-type material, but there are then very few holes in the valence band (because of the constancy of the np product). The opposite is true for a p-type semiconductor. Hence, the radiative recombination that occurs at thermal equilibrium produces very few photons, which are reabsorbed before leaving the crystal. In order to obtain significant light emission from a semiconductor, one must therefore somehow move away from thermal equilibrium to produce more electrons in the conduction band and more holes in the valence band.

The increased concentrations of holes and electrons can be created in several different ways, and the process is generally called *pumping* the material. For example, if high intensity, greater-than-bandgap light is directed onto the semiconductor, interband absorption occurs and results in many hole-electron pairs being created. These holes and electrons are initially *hot* but very rapidly *thermalize*, i.e., electrons settle to the bottom of the conduction band and holes rise to the top of the valence band through interaction with the lattice, as shown in Fig. 11.4. Thermalization is typically very fast, on the order of 10^{-14} s. After thermalization, holes and electrons recombine; this recombination can be radiative, resulting in the emission of a photon of approximately bandgap energy. The lifetime, or time it takes on the average for hole-electron pair to recombine after it is created, is typically about 10^{-11} s in a pumped semiconductor (with direct bandgap). Note that, in an indirect bandgap

Fig. 11.4 a, b

Thermalization of energetic carriers. After being pumped to higher energy levels, both electrons and holes settle to their respective bandedge states, **a** Direct bandgap material; **b** indirect bandgap material



semiconductor, thermalization causes holes and electrons to settle into states with significantly different wavevectors (Fig. 11.4). This means that a phonon is needed for recombination to be possible, as well as a hole and electron; hence it is a lower probability process. The typical lifetime for an indirect recombination is as much as 0.25 s; before that time has passed, the hole and electron on the average will have already recombined through some nonradiative process, giving up energy to the lattice or to defects, etc. Hence the efficiency of indirect radiative recombination is very low. The quantum efficiency for a direct bandgap material can approach 1, but for an indirect gap it is generally 0.0001 or lower. However, indirect emissive transitions do occur, and in some indirect bandgap materials (for example, GaP) measurable and indeed usable light emission does occur. Nevertheless it is extremely difficult to make lasers in indirect gap materials.

The optical pumping method that has been described as a means of producing the required increased concentration of electrons and holes is somewhat unwieldy in that a high intensity source is required; the process is inefficient, generating a great deal of heat. Also, the output must be optically filtered to separate the semiconductor emission from reflected pump light.

An alternative pumping technique, which is both efficient and extremely simple, employs the properties of a p-n junction. The energy band diagram of a p-n junction light emitter is shown in Fig. 11.5. Under zero bias conditions, there are many electrons in the conduction band on the n-side and many holes in the valence band on the p-side, but very few can get over the barrier and into the junction region. When a forward bias voltage V_0 is applied, the barrier is lowered and many holes and electrons are injected into the junction region, where they recombine, resulting in the generation of photons. The light emitted under stimulation from a battery is called electroluminescence, while that produced by a pumping with a shorter wavelength optical source is called photoluminescence. A typical electroluminescent diode, or light-emitting diode (LED), is diagrammed in Fig. 11.6. Light is generated in the p-n junction and leaves the diode by passing through bulk material outside of the junction region. Much of the light is reabsorbed before it leaves the diode. Hence one speaks of external quantum efficiency, defined by

$$\eta_{\text{ext}} = - \frac{\text{number of photons emitted in desired direction}}{\text{number of hole - electron pairs injected}} \quad (11.18)$$

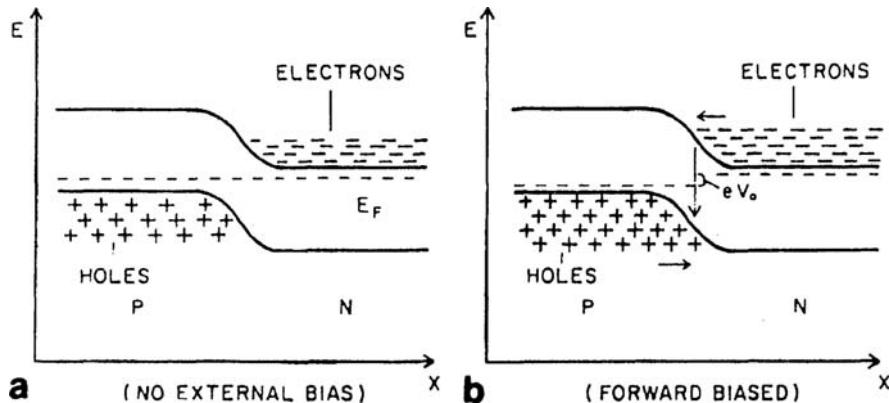


Fig. 11.5 a,b Energy band diagram of a p-n junction light emitter; **a** no external bias voltage applied; **b** external bias voltage V_0 applied with forward polarity

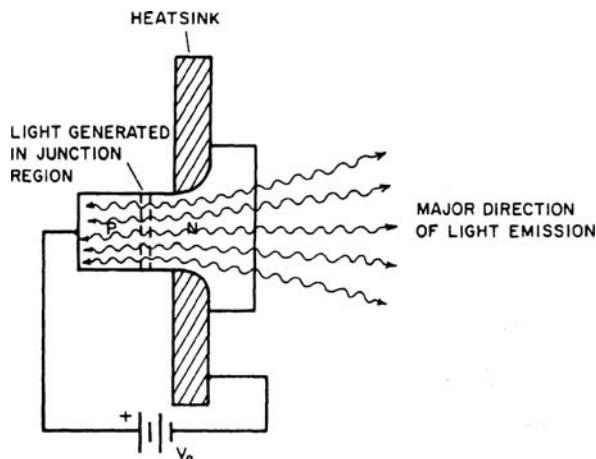


Fig. 11.6 A p-n junction light-emitting diode (LED)

as well as internal quantum efficiency, given by

$$\eta_{\text{int}} = \frac{\text{number of photons generated}}{\text{number of hole - electron pairs injected}} \quad (11.19)$$

In designing a light emitter, one would naturally try to minimize reabsorption by making the layer of material between the junction and the surface very thin, and by choosing a material with a very small absorption coefficient if possible. Such diodes have obtained wide usage as panel indicators and display devices because they are small, cheap, more reliable than incandescent devices, and easily incorporated into integrated circuit arrays. LED's are also very useful as light sources for fiber optic

communications because they can be modulated at high frequencies (although not as high as for lasers) and can be conveniently coupled to the cores of multimode fibers.

The wavelength spectrum of semiconductor light emitters, regardless of the pumping technique used, is relatively simple when compared, for example, to that of gaseous discharge sources, which typically have numerous lines or emission peaks. A good semiconductor light emitter with low defect or extraneous impurity content has a single emission peak at approximately the band-edge wavelength, corresponding to the energy gap. This peak has a halfwidth of typically 200–300 Å, which is considerably wider than gaseous source lines. Extraneous impurities and defect states within the bandgap allow radiative transitions which produce secondary emission peaks at other wavelengths. These secondary peaks indicate that some of the injected hole-electron pairs recombine to produce photons in other than the major band-edge peak. Hence the quantum efficiency at the desired wavelength is reduced. A thorough discussion of the emission spectrum of semiconductor light emitters is beyond the scope of this book. However relatively complete data are available in the literature for most of the commonly used semiconductors [5].

Even if a semiconductor light emitter is formed in relatively pure and defect free material, so that the light emission is almost entirely at the band-edge wavelength, the peak wavelength is still dependent on temperature and doping, changes the effective bandgap by allowing emissive transitions between donor and acceptor states. Thus the emission wavelength of a semiconductor light emitter can be varied over a somewhat limited range by doping it appropriately. For example, consider the case of GaAs doped with the n-type dopants Te, Se or Sn. The center wavelength of the emission peak changes as shown in Fig. 11.7. At relatively low doping levels, recombination occurs between electrons filling donor states 6 meV below conduction band edge and empty states at top of valence band. As doping increases above $5 \times 10^{17} \text{ cm}^{-3}$, the donor band merges with the conduction band and recombination occurs between filled conduction band states near the quasi fermi level and empty valence band states. The p-type dopants Cd and Zn also cause a shift in the emission peak wavelength, as shown in Fig. 11.8. At low p-type dopant concentration, the recombination is from filled states at the bottom of the conduction band to

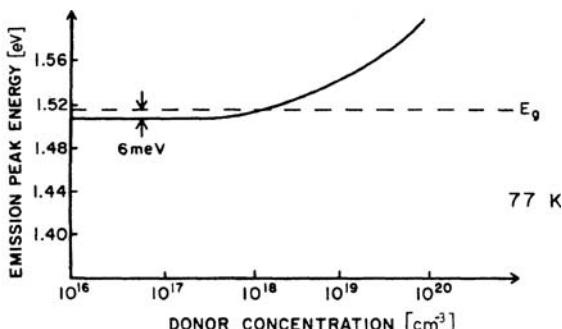
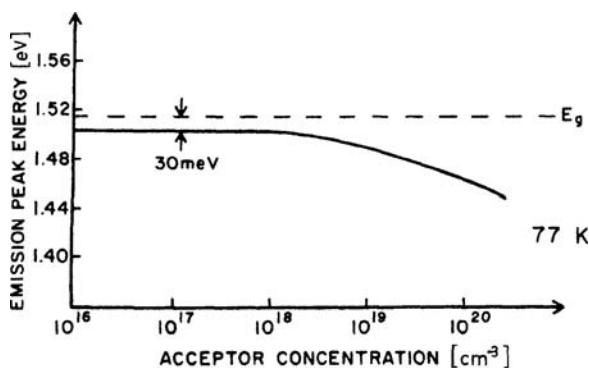


Fig. 11.7 Peak emission wavelength for n-type GaAs as a function of donor concentration [6]

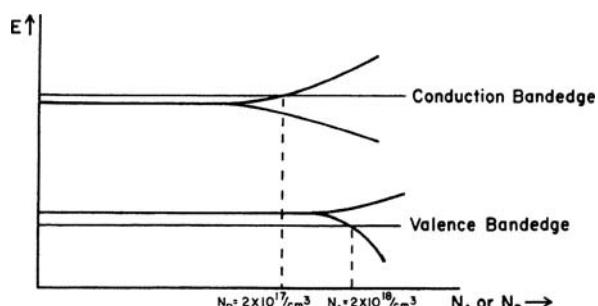
Fig. 11.8 Peak emission wavelength for p-type GaAs as a function of acceptor concentration [7]



empty acceptor states located 30 meV above the valence band edge. As the doping is increased above 10^{18} cm^{-3} the acceptor states broaden into a band, as shown in Fig. 11.9 and recombination occurs between filled states at the bottom of the conduction band and empty states at the top of the acceptor band. Thus the emission peak energy decreases for p-type doping, in contrast to the way it increases for n-type. As the p-type doping is increased above about $2 \times 10^{18} \text{ cm}^{-3}$, the acceptor band merges with the valence band, and hence the radiative recombination may be described as band-band, although it is really from conduction band states to acceptor states. The recombination does not occur only between states at exactly the same energy in the conduction and valence bands, respectively. Instead it occurs between states spread over a range of energies. Hence the emission is not single wavelength but rather has a line-width on the order of hundreds of Angstroms. As the doping is increased for either n- or p-type GaAs, the line-width of emission increases from typically 100 Å, for lightly doped material, to several hundred Angstroms, for heavily doped material. This increase in line-width is due to impurity banding, which makes available pairs of filled and empty states for recombination over a range of energies.

In summary, the principal properties of light emitted by the spontaneous recombination of electrons and holes in a semiconductor are that photon wavelengths

Fig. 11.9 Donor and acceptor energy levels (bands) in GaAs as a function of dopant concentration [8]



are spread over a range of several hundred Angstroms, with the peak wavelength corresponding approximately to the band-edge, but also depending on dopant concentration and type. Photons are emitted more-or-less isotropically, except for the effects of emitter geometry, and there is no fixed phase relationship between photons. The generation of each photon is a spontaneous event, completely unrelated to any other photon. The characteristics of light generated by stimulated emission contrast strongly with those of spontaneously emitted light.

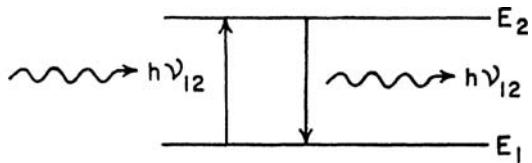
11.2.2 Stimulated Emission

Spontaneously emitted photons can, of course, be reabsorbed by any of the absorption processes that have been previously described. Alternatively, however, if the photon happens to encounter an electron hole pair of proper energy separation, it can encourage the two to recombine; in this case, a new photon is produced which is an exact duplicate of the one that stimulated the recombination process, i.e., it has exactly the same energy (frequency), direction, phase and polarization. This is stimulated emission, which occurs in addition to spontaneous emission, so that in general the intensity of the radiation produced in the presence of external radiation consists of two parts. One part is independent of the external radiation (spontaneous emission); the other part, has intensity proportional to the external radiation, while its frequency, phase, direction of propagation and polarization are the same as those of the external radiation. Note that the *external radiation* can be that generated by other spontaneous and/or stimulated emissions within the sample, and does not have to be truly *external* field.

Except under very special conditions, stimulated emission is less probable than spontaneous emission, and hence a normal light emitter produces mostly spontaneous emission. One has to make a special effort to get a significant amount of stimulated emission. First, the concentrations of holes and electrons in the light emitting region must be so large that most of the states near the bottom of the conduction band are filled with electrons and most of the states near the top of the valence band are empty. This makes it more probable for the photon to cause the transition of an electron from the conduction band to the valence band, producing a duplicate photon, than to be absorbed by causing an electron to move from the valence band up to an empty state in the conduction band. Secondly, in order to have stimulated emission predominate over spontaneous emission, there must be an intense field of photons of the proper energy. The necessity of satisfying these two conditions can be demonstrated mathematically by using a simple two energy level model [9], as shown in Fig. 11.10.

At thermal equilibrium (i.e. in the absence of any external *pumping* energy source), most of the electrons in the material will be in the lower energy level. The electron concentrations N_1 and N_2 in the two levels are related by the following expression from Boltzmann statistics:

Fig. 11.10 Simplified two energy level model for light emission



$$\frac{N_2}{N_1} = e^{-(E_2 - E_1)/kT}, \quad (11.20)$$

where it has been assumed that the two levels contain an equal number of available states. Consider the situation that exists when the material is immersed in a radiation field of photons with energy $h\nu_{12}$, equals to $E_2 - E_1$, such that the total energy in the field per unit volume (at the frequency ν_{12}) is $\rho(\nu_{12})$. The processes that can then occur are absorption, spontaneous emission and stimulated emission. For the steady-state case, these processes are linked by a transition rate equation given by, [9]

$$\underbrace{B_{12}N_1\rho(\nu_{12})}_{\text{absorption}} = \underbrace{A_{21}N_2}_{\text{spontaneous emission}} + \underbrace{B_{21}N_2\rho(\nu_{12})}_{\text{stimulated emission}} \quad (11.21)$$

where B_{12} , B_{21} and A_{21} are proportionality constants known as Einstein coefficients. Note in (11.21) that upward or downward electronic transitions are always proportional to the number of electrons in the initial level, N_1 , or N_2 , respectively, and that both absorption and stimulated emission are proportional to the optical field strength $\rho(\nu_{12})$ as well. The conditions necessary to produce predominance of stimulated emission over both absorption and spontaneous emission can be seen by taking the ratios of the respective transition rates, from (11.21) given by

$$\frac{\text{stimulated emission rate}}{\text{absorption rate}} = \frac{B_{21}N_2\rho(\nu_{12})}{B_{12}N_1\rho(\nu_{12})} = \frac{B_{21}}{B_{12}} \frac{N_2}{N_1}, \quad (11.22)$$

$$\frac{\text{stimulated emission rate}}{\text{spontaneous emission rate}} = \frac{B_{21}N_2\rho(\nu_{12})}{A_{21}N_2} = \frac{B_{21}}{A_{21}}\rho(\nu_{12}). \quad (11.23)$$

It can be seen from (11.20) that N_2 is much less than N_1 , at thermal equilibrium, for the usual case of $E_2 - E_1$ being much greater than kT . Hence, for stimulated emission to predominate over absorption, the ratio N_2/N_1 in (11.22) must be greatly increased by introducing a pumping source of energy, as described in Section 11.2.1. The condition of N_2 being greater than N_1 , is called an inverted population. From (11.23) it is obvious that an intense flux of photons of the proper energy, i.e. large $\rho(\nu_{12})$, is required in order to make stimulated emission much greater than spontaneous emission. This photon concentration is usually produced by introducing positive optical feedback from mirrors or other reflectors.

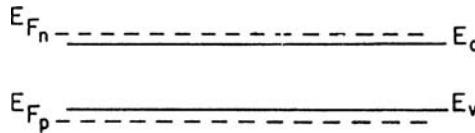


Fig. 11.11 Energy band diagram of a semiconductor with external energy input. E_{Fn} and E_{Fp} are the quasi-Fermi levels for electrons and holes respectively

The two requirements that there be an inverted population, and that an intense flux of photons of the proper energy be present, are common to all types of lasers. The simplified two-level model that has been used to demonstrate these principles is more appropriate for a gas laser than for semiconductor laser, because the electrons and holes in a semiconductor are distributed over bands of energy rather than occupying two discrete energy levels. Nevertheless the same principles apply. In a semiconductor that is undergoing pumping, electrons fill states at the bottom of the conduction band roughly up to an energy level E_{Fn} , while holes fill states at the top of the valence band approximately down to an energy level E_{Fp} , as shown in Fig. 11.11. The energy levels E_{Fn} and E_{Fp} are known as the quasi-Fermi levels for electrons and holes [10], since strictly speaking, the Fermi level is defined only for the case of thermal equilibrium. For the nonequilibrium case of external energy input, quasi-Fermi levels must be defined separately for holes and electrons. From Fig. 11.11, it can be seen that a state of population inversion is experienced by any photon of energy $h\nu_{\text{phot}}$ when the condition

$$E_{Fn} - E_{Fp} > h\nu_{\text{phot}} \quad (11.24)$$

is satisfied. The minimum energy condition for population inversion to exist is, of course, given by

$$E_{Fn} - E_{Fp} = E_c - E_v = E_g, \quad (11.25)$$

where E_g is the bandgap energy. Thus, in a pumped semiconductor, population inversion exists, and hence stimulated emission dominates over absorption for photons with energy in the range given by

$$E_g \leq h\nu_{\text{phot}} \leq (E_{Fn} - E_{Fp}). \quad (11.26)$$

11.3 Lasing

When the two conditions of population inversion and high photon density are satisfied, a selection process takes place which alters the character of the radiation and

results in the generation of significant amounts of stimulated emission. If a suitable optically resonant structure is provided, lasing can occur.

11.3.1 Semiconductor Laser Structures

The conventional structure used for a discrete semiconductor laser is a rectangular parallelepiped such as that shown in Fig. 11.12, which is produced by cleaving the crystal along lattice planes that are perpendicular to the light emitting layer. For example a (100) wafer of GaAs can be easily cleaved along the four perpendicular (110) planes to produce such a structure. Typical dimensions for the length L and width W are $300 \mu\text{m}$ and $50 \mu\text{m}$, respectively, but great variations from the values are possible, as is discussed in detail in Chapter 12. The two end faces of width W form an optically resonant, reflective structure known as a Fabry-Perot etalon. When such a structure is provided, increasing the pumping power will result in lasing when a certain critical threshold level is exceeded.

11.3.2 Lasing Threshold

As pumping power is increased from zero, an increasing number of photons are emitted by spontaneous emission. Even when pumping power exceeds that required to produce population inversion, the emission, at first, is still composed of predominantly spontaneously generated photons. The spontaneously emitted photons are produced more-or-less equally in all directions, as shown in Fig. 11.12. However, most of these photons start off in directions that very soon carry them out of the

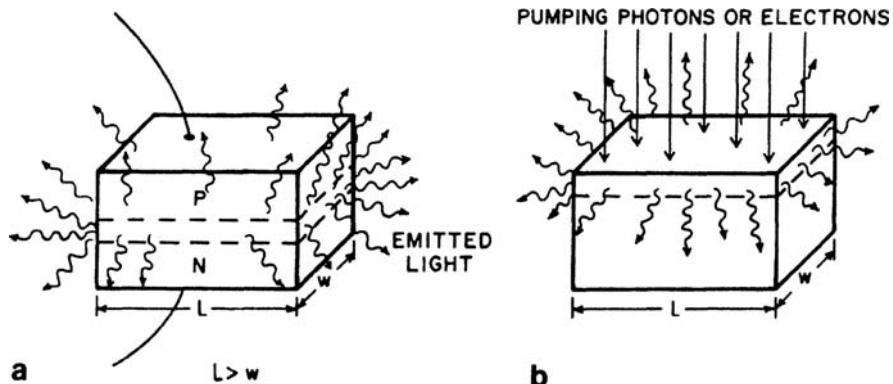


Fig. 11.12 a,b Semiconductor laser structure; **a** p-n junction pumped (laser diode); **b** electron or photon beam pumped. Light emission is shown as it would be for pumping levels below the threshold for lasing. (See text for explanation). The dashed lines mark the extremes of the pumped region

inverted population region where predominant stimulated emission can occur, and thus they are unable to reproduce themselves. The few that happen to be traveling in the plane of the inverted population region are able to duplicate themselves by stimulated emission many times before they emerge from the laser. In addition, for any given energy bandgap and distribution of holes and electrons, there will be one particular energy (wavelength) which is preferred (transitions more probable) over others. This wavelength or energy usually corresponds to first order to the peak wavelength at which spontaneous emission takes place in the material. As a result of this preferred energy and direction, when stimulated emission builds up, the emitted radiation narrows both in spatial divergence and spectral linewidth (reduced from several hundred Angstroms to about 25 Å). Note that, as stimulated emission builds up, more hole-electron pairs are consumed per second. Hence, for any given input hole-electron pair generation rate, spontaneous emission is suppressed, because the stimulated emission consumes the generated pairs before they can recombine spontaneously. Such predominant stimulated emission as has been described is known as *super-radiance* radiation. It is not yet the coherent light produced by lasing. Such radiation is due to amplified spontaneous radiation. The photons are not traveling in *exactly* the same direction; they may have slightly different energy, they are not in phase. In other words, the radiation is still incoherent.

When the light emitting region is bounded by parallel reflecting planes as in Fig. 11.12, some reflection will occur at the surfaces and therefore some of the radiation will pass back and forth through the inverted population region and be amplified further. In order to be reflected back and forth in the narrow inverted population layer photons must be *exactly* in the plane of the layer and *exactly* perpendicular to the Fabry-Perot surfaces. The radiation must also be of uniform frequency and phase to avoid destructive interference, because the Fabry-Perot reflectors form a resonant *cavity* in which optical modes are established. Thus, only that small fraction of the super-radiance radiation [11] that fulfills all of these requirements continues to survive and becomes the dominant species of radiation which is produced. Such radiation, uniform in wavelength, phase, direction and polarization is coherent radiation. The device is said to be lasing.

The transition between the super-radiance region of output and the lasing region of output (as input power is increased) is taken to be the point at which the amplification of the preferred mode by stimulated emission becomes greater than absorption and all other photon losses, so that the radiation bounces back and forth between Fabry-Perot surfaces, growing in intensity with each bounce (i.e., oscillation sets in). Of course, the radiation does not continue to build up in intensity indefinitely. As the intensity grows, it stimulates more recombinations until concentrations of holes and electrons are being used up at exactly the same rate at which they are being generated by the pumping source, and a dynamic equilibrium is established. In this steady state condition, the laser oscillates at a single peak wavelength (or possibly several discrete wavelengths corresponding to different longitudinal modes) which is highly monochromatic (approximately 1 Å linewidth). A standing wave is produced within the crystal, with an integral number of half-wavelengths existing between the Fabry-Perot faces for each mode. Because of the previously described

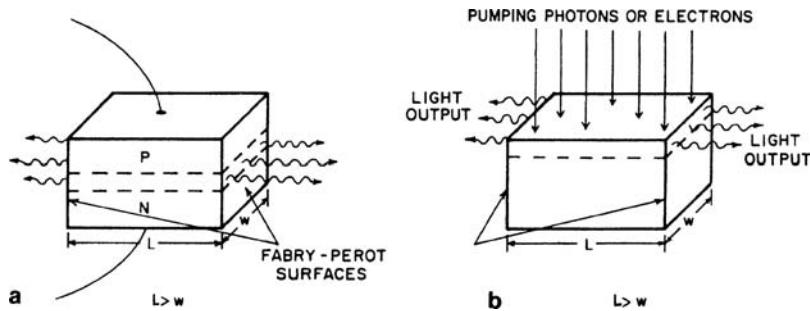


Fig. 11.13 a,b Semiconductor laser operating above the threshold for lasing. **a** p-n junction pumped; **b** electron or photon beam pumped

directional selectivity of the reflecting and faces, photons of the lasing mode(s) are emitted predominantly in the direction normal to the end faces, as shown in Fig. 11.13. The light waves (photons) emitted from such a resonant structure are in phase, and in addition, the waves which may be most easily propagated are those with their electric vector perpendicular to the plane of the inverted population layer. Hence the emitted light is normally polarized in that manner.

The transitions between the regions of spontaneous emission, super-radiance emission and laser emission, as the pumping power is increased, are very abrupt. In fact, although the transition between super-radiance and lasing has been described in detail for pedagogic purposes, the actual transition in a normal laser is so abrupt that super-radiance emission cannot be observed. As the pumping power goes above the threshold point, the character of the light emitted abruptly changes to the lasing type, with linewidth decreasing from several hundred Angstroms to one Angstrom, and phase coherence being observable. The intermediate stage of narrowed linewidth due to predominant stimulated emission without phase coherence (i.e. super-radiance) is passed through so quickly that observations cannot be made. Only in laser diode structures with damaged Fabry-Perot reflectors that inhibit lasing, can super-radiance emission be observed.

11.3.3 Efficiency of Light Emission

One important characteristic of semiconductor laser light emission, which has not yet been discussed, is the efficiency of conversion of electrical energy into optical energy. A laser diode operating above threshold is about a hundred times more efficient than a light emitting diode made in the same material. This improved efficiency results from several factors.

As previously mentioned, generation of a photon by stimulated emission has a relatively low probability of occurrence compared to that of spontaneous emission in the case of a noninverted population, such as exists in a light emitting diode. However, when the population is inverted, as in a laser, the probability of stimulated

emission becomes very large, even larger than the probabilities of either spontaneous emission or nonradiative recombination of the hole-electron pair. Thus the internal quantum efficiency is about a factor of ten higher in a laser diode than in an LED. In fact, if a laser diode is cooled to 77 K to reduce the thermal spread of electron and hole energies, the internal quantum efficiency can approach 100%.

Once the photons are generated, they also experience less loss in a laser than in an LED. Photons in the lasing mode travel mostly in the inverted population region, in which inter-band absorption is suppressed by a lack of electrons near the valence band edge and empty states at the conduction band edge. Photons in an LED travel mostly in noninverted, bulk material. Thus internal reabsorption of generated photons is much less in a laser. Also, the photons emitted from a laser travel in a much less divergent beam than those emitted from an LED, making it possible to collect (or couple) more of the light output at a desired location. Because of the combined effects of reduced internal reabsorption, greater beam collimation and increased internal quantum efficiency, a semiconductor laser can be expected to have an external quantum efficiency roughly a hundred times greater than its LED counterpart. Since both the laser and LED operate at approximately the same voltage, an increase in external quantum efficiency translates into a proportional increase in overall power efficiency.

In this chapter, the basic principles of light emission in semiconductor lasers and LED's have been discussed, and a distinction has been made between the characteristics of the light emitted by each. In Chapter 12, the laser diode is considered in more detail, and quantitative relations are developed to evaluate such important characteristics as mode spacing, threshold current density and efficiency. In addition, some other basic types of semiconductor lasers are discussed.

Problems

- 11.1(a) Why do semiconductor light emitting diodes have a characteristic emission line width on the order of hundreds of Angstroms while semiconductor lasers have line widths $\sim 1 \text{ \AA}$?
- (b) In order to produce a semiconductor laser, what two conditions basic to the gain mechanism must be satisfied?
- (c) To what extent is the light output of a semiconductor laser collimated? (Give a typical divergence angle.) What can be done to produce greater collimation?

11.2 Given that

$$B_{12}N_1\rho(v_{12}) = A_{21}N_2 + B_{21}N_2\rho(v_{12})$$

show that for a very large photon density $\rho(v_{12}) \rightarrow \infty$ it follows that

$$B_{12} = B_{21}.$$

- 11.3 If the absorption of a photon causes an electron to make a direct transition (i.e., no phonon involved) from a state with wavevector $|k| = 1.00 \times 10^7 \text{ cm}^{-1}$ in a given direction to a state with wavevector $|k| = 1.01 \times 10^7 \text{ cm}^{-1}$ in the same direction, what was the wavelength of the photon?
- 11.4 Why is a semiconductor laser inherently more efficient than a light-emitting-diode?
- 11.5 Sketch three different plots of typical diode laser emission spectra which show output light intensity vs. wavelength a) below threshold, b) at threshold, c) above threshold.
- 11.6 (a) Draw the (E vs. x) energy band diagram of a $p - n$ junction laser under forward bias and indicate the inverted population region, as well as E_{Fn} , E_{Fp} , E_c , E_v and E_g .
 (b) What energies establish the long and short wavelength limits of the laser emission spectrum?
- 11.7 Why is a semiconductor laser inherently capable of producing a higher optical power density (i.e. watts/ cm³) than is a gas laser?
- 11.8 Why are the Fermi levels in a semiconductor laser that is being pumped with external energy called “quasi-Fermi levels”?
- 11.9 Describe the characteristics of the light from a semiconductor laser diode as compared to those of a semiconductor LED.
- 11.10 In GaAs, if an electron makes a direct radiative transition from a state of energy = 20 meV above the conduction band edge to a state of energy = 35 meV below the valence band edge, what is the wavelength of the emitted photon?

References

1. O. Madelung: *Introduction to Solid-State Theory*, Springer Ser. Solid-State Sci., Vol. 2 (Springer, Berlin, Heidelberg third printing 1996) p. 57
2. E.L. Hill: The theory of relativity, in *Handbook of Physics*, E.U. Condon, H. Odishaw (eds.) (McGraw-Hill, New York 1967) pp. 2–44
3. E.W. Williams, H.B. Bebb: Gallium arsenide photoluminescence, in *Transport and Optical Phenomena*, (eds.) R.K. Williardson, A.C. Beer (eds.), Semiconductor and Semimetals. Vol. 8 (Academic, New York 1972) Chap. 5
4. J.I. Pankove (ed.): *Electroluminescence*, Topics Appl. Phys. Phys., Vol. 17 (Springer, Berlin, Heidelberg 1977) Chap. 3
5. J.I. Pankove (ed.): *Display Devices*, Topics Appl. Phys., Vol. 40 (Springer, Berlin, Heidelberg 1980) Chap. 2
6. D.A. Cusano: Solid State Commun. **2**, 353 (1964)
7. D.A. Cusano: Appl. Phys. Lett. **7**, 151 (1964)
8. G. Lucovsky, A. Varga: J. Appl. Phys. **35**, 3419 (1964)

9. B.G. Streetman: *Solid State Electronics*, 4th edn. (Prentice-Hall, Englewood Cliffs, NJ 1995) pp. 377–380
10. M. Shur: *Physics of Semiconductor Devices* (Prentice-Hall, Englewood Cliffs, NJ 1990) pp. 92–102
11. R. Bonifacio (ed.): *Dissipative Systems in Quantum Optics*, Topics Curr. Phys., Vol. 27 (Springer, Berlin, Heidelberg 1982)

Chapter 12

Semiconductor Lasers

In the preceding chapter, we discussed the basic principles of light emission in a semiconductor. Probably the most significant feature of this light emission is that it is possible to design a light source in such a way that stimulated emission of photons will predominate over both spontaneous emission and absorption. If a resonant reflecting structure such as a pair of plane, parallel end faces is provided, a lasing mode can be established and coherent optical emission will result. In this chapter, we consider several basic semiconductor laser structures and develop the quantitative theory necessary to calculate their expected performance characteristics.

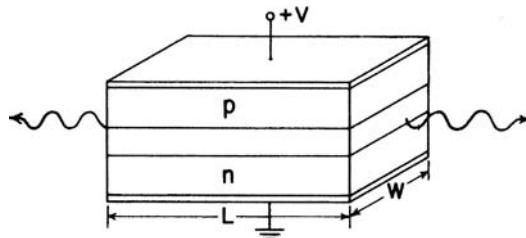
12.1 The Laser Diode

Since its inception in 1962 [1–4], the semiconductor laser diode has been developed from a laboratory curiosity into a reliable and marketable product, which has gained widespread acceptance as a source of light in many varied applications. In this section, we discuss the basic structure and performance of the discrete laser diode as opposed to the monolithically integrated laser diode of the optical integrated circuit, which will be covered in Chapters 14 and 15.

12.1.1 Basic Structure

The p-n junction laser diode is an excellent light source for use in optical integrated circuits and in fiber-optic signal transmission applications because of its small size, relatively simple construction and high reliability. To date, most laser diodes have been made in GaAs, $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$, or $\text{Ga}_x\text{In}_{(1-x)}\text{As}_{(1-y)}\text{P}_y$. However, fabrication technologies have been developed for producing lasers in other materials, such as GaN and SiGe, in order to obtain emission at different wavelengths. These newer lasers are described in Section 12.4. The basic structure of a p-n junction laser diode is shown in Fig. 12.1. A p-n junction is usually formed by epitaxial growth of a p-type layer on an n-type substrate. Ohmic contacts are made to each region to permit the flow of electrical current which is the *pumping* energy

Fig. 12.1 Basic structure of a p-n junction laser



source required to produce the inverted population in the active region adjacent to the junction. Two parallel end faces are fabricated to function as *mirrors* providing the optical feedback necessary for the establishment of a lasing optical mode. The device shown in Fig. 12.1 is a discrete diode structure, such as might be used in conjunction with an optical fiber transmission line. In later chapters, we shall consider a much more sophisticated, multi-layered structure with more elaborate optical feed-back schemes that are better suited for monolithic integration into optical integrated circuits. However, the fundamental structure provides a convenient basis for the development of a theoretical description of laser performance that can be easily adapted to more complex devices.

12.1.2 Optical Modes

The partially reflecting end faces of the laser diode provide an optical feedback that leads to the establishment of one or more longitudinal optical modes. Because of their similarity to the plane, parallel mirrors of a Fabry-Perot interferometer [5], the laser end faces are often called Fabry-Perot surfaces. When current is passed through the laser diode, light can be generated in the resulting inverted population laser by both spontaneous and stimulated emission of photons as explained in Chapter 11. Because of the reflection that occurs at the Fabry-Perot surfaces, some of the photons will pass back and forth many times through the inverted population region and be preferentially multiplied by stimulated emission. Those photons that are traveling exactly in the plane of the layer and exactly perpendicular to the Fabry-Perot surfaces have the highest probability of remaining in the inverted population layer where they can reproduce themselves by stimulated emission. Hence, they become the photons of the optical mode or modes that are established when steady state operation is achieved at a given current level. It is possible for other modes to develop so as to correspond to photons following a zig-zag path reflecting off of the side faces of the laser, but in a practical device these modes are usually suppressed by roughening the side faces or by using some equivalent technique to attenuate the undesired mode. The radiation of the lasing mode must also be of uniform frequency and phase to avoid destructive interference. As a result, a standing wave is produced within the laser diode with an integral number of half-wavelengths between the parallel faces.

The mode number m is given by the number of half-wavelengths. Thus

$$m = \frac{2Ln}{\lambda_0}, \quad (12.1)$$

where L is the length between end faces, n is the index of refraction of the laser material, and λ_0 is the vacuum wavelength of the emitted light. The mode spacing can be determined by taking $dm/d\lambda_0$, keeping in mind that semiconductor lasers are always operated near the bandgap wavelength, where n is a strong function of wavelength. Thus,

$$\frac{dm}{d\lambda_0} = -\frac{2Ln}{\lambda_0^2} + \frac{2L}{\lambda_0} \frac{dn}{d\lambda_0}. \quad (12.2)$$

For $dm = -1$, the mode spacing $d\lambda_0$ is given by

$$d\lambda_0 = \frac{\lambda_0^2}{2L \left(n - \lambda_0 \frac{dn}{d\lambda_0} \right)}. \quad (12.3)$$

We take dm equal to -1 because a decrease of one in the value of m corresponds to one less half-wavelength between the Fabry-Perot end faces, i.e., an increase in wavelength λ_0 .

A typical mode spectrum for a diode laser is shown in Fig. 12.2. Usually several longitudinal modes will coexist, having wavelengths near the peak wavelength for spontaneous emission. The mode spacing for a GaAs laser is typically $d\lambda_0 \cong 3 \text{ \AA}$. In order to achieve single mode operation, the laser structure must be modified so as to suppress all but the preferred mode. We will consider such specialized devices in later chapters.

12.1.3 Lasing Threshold Conditions

When a laser diode is forward biased and current begins to flow, the device does not immediately begin lasing. At low current levels, the light that is emitted is

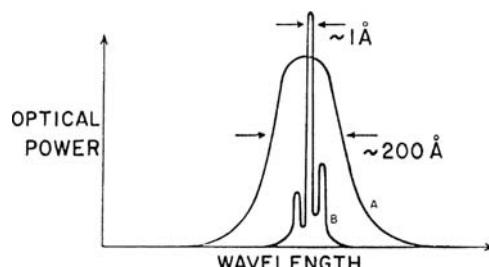


Fig. 12.2 Optical emission spectrum of the diode laser; A spontaneous emission below threshold; B lasing mode structure above threshold

mostly due to spontaneous emission and has a characteristic spectral linewidth on the order of hundreds of Angstroms. It is incoherent light. As the pumping current is increased, a greater population inversion is created in the junction region and more photons are emitted. The spontaneously emitted photons are produced going more-or-less equally in all directions. Most of these start off in directions that very soon carry them out of the inverted population region where net stimulated emission can occur, and thus are unable to reproduce themselves. However, those few photons that happen to be traveling exactly in the junction plane and perpendicular to the reflecting end faces are able to duplicate themselves many times before they emerge from the laser. In addition, for any given energy bandgap and distribution of holes and electrons, there is one particular energy(wavelength) that is preferred over others. To first order, this wavelength usually corresponds to the peak wavelength at which spontaneous emission takes place in the material. As a result of this preferred energy and direction, when stimulated emission builds up with increasing current, the emitted radiation narrows substantially both in spectral linewidth and spatial divergence. As stimulated emission builds up, the photon density (intensity) of the optical mode increases, thus leading to a further increase in stimulated emission so that more hole-electron pairs are consumed per second. Hence, spontaneous emission is suppressed for any given input hole-electron pair generation rate, because the stimulated emission uses up the generated pairs before they can recombine spontaneously. Because of the phase condition of (12.1), the light produced by stimulated emission in an optically resonant structure like that of the laser diode is coherent, and the device is said to be lasing.

The transition from nonlasing emission to lasing emission occurs abruptly as the current level exceeds the threshold value. As the threshold current is exceeded, the onset of lasing can be experimentally observed by noting the sharp break in the slope of the optical power versus pump current curve (Fig. 12.3), which results because of the higher quantum efficiency inherent in the lasing process (as explained in Section 11.3.3). Also, the spectral line-shape of the emitted light abruptly changes from the broad spontaneous emission curve to one consisting of a number of narrow modes as shown in Fig. 12.2. Quantitatively speaking, the lasing threshold corresponds to the point at which the increase in the number of lasing mode photons (per second) due to stimulated emission just equals the number of photons lost

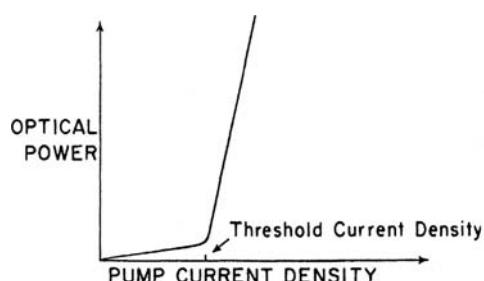


Fig. 12.3 Light output of diode laser as a function of input pumping current

(per second) because of scattering, absorption or emission from the laser. In conventional terms used to describe an oscillator one would say that the device has a closed loop gain equal to unity. Using this fact, it is possible to develop an expression for the threshold current as a function of various material and geometric parameters.

We will begin by considering the p-n junction laser structure shown in Fig. 12.1 and by following the approach used by Wade et al. [6]. The light is emitted from the laser preferentially in the direction perpendicular to the Fabry-Perot surfaces. The transverse spatial energy distribution of the light wave (photon density) is as shown in Fig. 12.4. The photon distribution extends or spreads into the inactive(noninverted) regions on each side of the junction due mainly to diffraction. Thus, there is a light emitting layer of thickness D , which is greater than the thickness d of the active or inverted population layer. For example, in GaAs diodes, $d \cong 1 \mu\text{m}$, $D \cong 10 \mu\text{m}$. It can be seen from the idealized spatial energy diagram that, of the total number of photons existing in the lasing mode at any given time, only a fraction d/D remain in the active region and can generate additional photons by stimulated emission. This effect reduces the gain available from the device.

To derive a quantitative expression giving the required current density to produce lasing, consider a single pass of a wave of laser light (photon flux) from one Fabry-Perot surface to the other. The power flow diagram looks like that given in Fig. 12.5. P_s is the optical power incident internally on each end face and R is the power reflection coefficient. Under oscillating (lasing) conditions, RP_s grows exponentially with distance during a single pass reaching the value P_s at the opposite Fabry-Perot (F-P) surface. Loss mechanisms present are overcome by the laser gain mechanism of stimulated emission. At each F-P surface, power of $\frac{1}{2} P_{\text{out}} = (1 - R)P_s$ is emitted. If α is the loss coefficient [cm^{-1}] for the wave as it travels (including all types of optical loss) and g is the gain coefficient [cm^{-1}], the power as a function of distance is given by

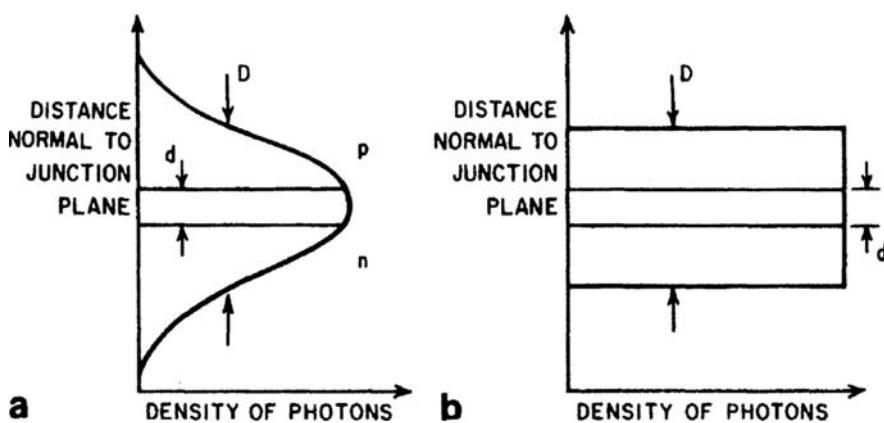
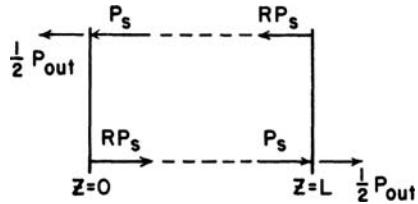


Fig. 12.4 a,b Transverse spatial energy distribution of the light in a diode laser: **a** actual; **b** idealized (area under the idealized curve is adjusted so that d/D remains unchanged)

Fig. 12.5 Power flow diagram for a diode laser



$$P = RP_s \exp\left(g \frac{d}{D} - \alpha\right) Z. \quad (12.4)$$

Note that gain occurs only over the inverted population region; hence, g must be multiplied by d/D , while loss occurs wherever the field extends and hence α is not multiplied by anything. For oscillation (the loop gain equals 1), it must be true that

$$P_s = RP_s \exp\left(g \frac{d}{D} - \alpha\right) L. \quad (12.5)$$

or

$$\ln \frac{1}{R} = \left(g \frac{d}{D} - \alpha\right) L. \quad (12.6)$$

Thus,

$$g \frac{d}{D} = \alpha + \frac{1}{L} \ln \frac{1}{R}. \quad (12.7)$$

The gain coefficient g is related to the injected current density of holes and electrons. It can be shown that this relationship is [7]

$$g = \frac{\eta_q \lambda_0^2 J}{8\pi e n^2 d \Delta v}, \quad (12.8)$$

for GaAs at 300 K		
η_q :	internal quantum efficiency	0.7
λ_0 :	vacuum wavelength emitted	9×10^{-5} cm
n :	index of refraction at λ_0	3.34
Δv :	linewidth of spontaneous emission	1.5×10^{13} s ⁻¹
e :	electronic charge	
d :	thickness of active region	10^{-4} cm
J :	injected current density	

Thus, substituting (12.8), which holds both above and below threshold, into the threshold condition (12.7), one obtains

$$\overbrace{\frac{\eta_q \lambda_0^2 J_{\text{TH}}}{8\pi en^2 \Delta\nu D}}^{\text{effective gain per pass}} = \overbrace{\alpha + \frac{1}{L} \ln \frac{1}{R}}^{\text{loss per pass}}, \quad (12.9)$$

or

$$J_{\text{TH}} = \frac{8\pi en^2 \Delta\nu D}{\eta_q \lambda_0^2} \left(\alpha + \frac{1}{L} \ln \frac{1}{R} \right). \quad (12.10)$$

Thus, the threshold current (density) is that which is necessary to produce just enough gain to overcome the losses.

Note that, from the standpoint of threshold, the light output of the laser at the end faces must be counted as a loss. This is taken account by the term $(1/L)\ln(1/R)$. It may not be immediately obvious that this term represents *loss* due to the power emitted from the end faces but, if we substitute the transmission coefficient, $T = 1 - R$ and do a series expansion of $\ln[l/(1-T)]$, we get

$$\frac{1}{L} \ln \left(\frac{1}{R} \right) = \frac{1}{L} \ln \left(\frac{1}{1-T} \right) = \frac{1}{L} \left(T - \frac{T^2}{2} + \frac{T^3}{3} - \frac{T^4}{4} + \dots \right). \quad (12.11)$$

Discarding the higher-order terms in T one obtains

$$\frac{1}{L} \ln \left(\frac{1}{R} \right) \cong \frac{T}{L}. \quad (12.12)$$

T/L represents a loss coefficient [cm^{-1}] obtained by averaging the end *loss* coefficient T over the length. Since T has a value of typically 0.6 (for GaAs), discarding the T^2 and higher-order terms does not produce an accurate result and should not be done in quantitative work. However, the purpose of this example is to qualitatively demonstrate that the nature of the term $(1/L)\ln(1/R)$ is that of an *average* loss coefficient per unit length resulting from transmission of photons out of the laser.

From the theoretical development leading to (12.10), we see that the spreading of photons (optical fields) out of the inverted population region into surrounding passive regions results in a substantial increase in threshold current density. This fact suggests that the laser diode should be designed so as to make the ratio $D/d = 1$ for optimum performance. In Section 12.2 and in Chapter 14 we will discuss ways of making such *confined-field* lasers.

12.1.4 Output Power and Efficiency

Expressions for the overall power efficiency and the power output of a diode laser can be derived as follows: First, consider the losses over a small (incremental) distance ΔZ . To first order, the power loss over this distance is

$$\begin{aligned} P_{\text{loss}} &= P - Pe^{-\alpha \Delta Z} = P(1 - e^{-\alpha \Delta Z}) \\ &\cong P[1 - (1 - \alpha \Delta Z)] = \alpha P \Delta Z, \end{aligned} \quad (12.13)$$

or

$$dP_{\text{loss}} = \alpha P dZ. \quad (12.14)$$

Then the power absorbed over length L is

$$P_{\text{loss}} = \int_{Z=0}^{Z=L} dP_{\text{loss}} = \alpha \int_0^L P dZ. \quad (12.15)$$

Similarly, the power generated per pass is given by

$$P_{\text{gen}} = \int_{Z=0}^{Z=L} dP_{\text{gen}} = g \frac{d}{D} \int_0^L P dZ. \quad (12.16)$$

Thus, the internal power efficiency η is given by

$$\eta \equiv \frac{P_{\text{gen}} - P_{\text{loss}}}{P_{\text{gen}}} = \frac{g \frac{d}{D} - \alpha}{g \frac{d}{D}}. \quad (12.17)$$

Substituting from (12.7)

$$\eta = \frac{\frac{1}{L} \ln \frac{1}{R}}{\alpha + \frac{1}{L} \ln \frac{1}{R}}. \quad (12.18)$$

Note again that α is the total loss coefficient including both interband and free carrier absorption, as well as any scattering loss. The power out is then given by

$$P_{\text{out}} = \eta P_{\text{in}} = \eta \left[\frac{J}{e} \eta_q (L \times W) h\nu \right], \quad (12.19)$$

where P_{in} is the optical power generated within the laser, and L being the length, W the width and ν the lasing frequency. Substituting from (12.18) one obtains

$$P_{\text{out}} = \eta = \frac{\frac{1}{L} \ln \frac{1}{R}}{\alpha + \frac{1}{L} \ln \frac{1}{R}} \frac{J \eta_q}{e} (L \times W) h\nu. \quad (12.20)$$

(Note: P_{out} is power out of both end faces.) The overall power efficiency including the effect of series resistance in the device is thus given by

$$\eta_{\text{tot}} = \frac{P_{\text{out}}}{P_{\text{in,tot}}} = \frac{\frac{1}{L} \ln \frac{1}{R} \frac{J \eta_q}{e} (L \times W) h\nu}{\alpha + \frac{1}{L} \ln \frac{1}{R} \underbrace{\frac{J}{e} (L \times W) h\nu + [J(L \times W)]^2 R_{\text{series}}}_{l^2 R \text{ loss}}}. \quad (12.21)$$

In cases where the series resistance of the diode can be neglected this reduces to

$$\eta_{\text{tot}} = \frac{\frac{1}{L} \ln \frac{1}{R}}{\alpha + \frac{1}{L} \ln \frac{1}{R}} \eta_q. \quad (12.22)$$

Note that these power and efficiency formulae hold at and above threshold. For current densities below threshold (i.e., before oscillation has set in), the replacement $g(d/D) = \alpha + (1/L)\ln(1/R)$ does not hold, and the expressions would have to be left in terms of g , D , d and α rather than L , R and α .

The efficiency of semiconductor lasers decreases with increasing temperature because of two effects. First, absorption increases [resulting in a larger α in (12.18)] and secondly, the quantum efficiency (η_q) is reduced. This latter effect occurs because the holes and electrons are *smeared out* in energy over a wider range due to thermal excitation, and hence less hole-electron pairs of the proper energy separation are available for stimulated emission into the lasing mode at any given level of injected input current. These same two effects are responsible for the increase in threshold current density required for lasing, sec (12.10). Obviously, for a given input current density, the power out is reduced by increasing temperature due to these same two factors. Because of the afore-mentioned bad effects of heat, p-n junction lasers are usually operated on a pulse basis. The *turn-on* times and decay times of the diodes are very short ($\sim 10^{-10}$ s or shorter); hence, the diodes can easily be pulsed in good *square wave* 100 ns duration pulses. Overall power efficiencies are generally low (about a few percent), and peak power out is generally about 10 W. Semiconductor p-n junction lasers can be operated at room temperature (or higher) on a cw basis except that enough pumping current must be supplied to achieve threshold, and this much current must be less than that which would destroy the device due to heating. Typical threshold current densities for basic semiconductor p-n junction lasers like that of Fig. 12.1 are on the order of 10^4 A/cm², but since areas are about 10^{-3} cm², the peak current are typically on the order of 10 A. However, efficiency and power out are strong functions of material properties and device geometry. In later chapters, we will consider advanced geometries that will permit cw operation of laser diodes with threshold currents less than 20 mA.

12.2 The Tunnel-Injection Laser

We include within this chapter on basic types of semiconductor lasers the tunnel-injection laser, proposed by Wade et al. in 1964 [6, 8]. As the name implies, this device is pumped by a current of electrons or holes which reach the active region by *tunneling* through an energy barrier. The tunnel-injection laser was one of the first of the confined-field type lasers to be proposed [8–10], and it serves as the basis for a simple model that explains how field-confinement like that used in heterojunction lasers [11–13] results in substantially reduced threshold current density and higher efficiency. Modern heterojunction lasers, which are discussed in Chapter 14, have reached a high level of sophistication, providing single mode CW operation at room temperature.

12.2.1 Basic Structure

A semiconductor laser that in principle combines the best features of the p-n junction laser (small size, simplicity, low voltage power supply) within a device structure fabricated in a single type of semiconductor material, so that no junction formation required, is the tunnel-injection laser, shown in Fig. 12.6. In this laser, a single crystal of uniformly doped semiconductor material is used, without a junction. The hole-electron pairs are injected into the semiconductor by tunneling and diffusion. If a p-type semiconductor is used, electrons are injected through the insulator by tunneling while holes diffuse from the source at the contact of metal No. 1 to the semiconductor. If an n-type semiconductor is used, holes are tunneled through the insulator while electrons are injected at the metal No. 1 contact.

The process of tunneling can be better understood by referring to the energy band diagram of Fig. 12.7. The diagram shows the energy band picture after the required bias voltage has been applied and, as a result, the Fermi levels have shifted to new positions. Note that most of the voltage drops across the insulator; hence, a very

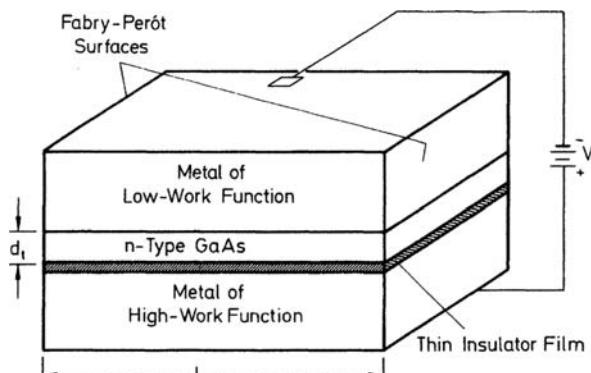


Fig. 12.6 Basic structure of a tunnel-injection laser

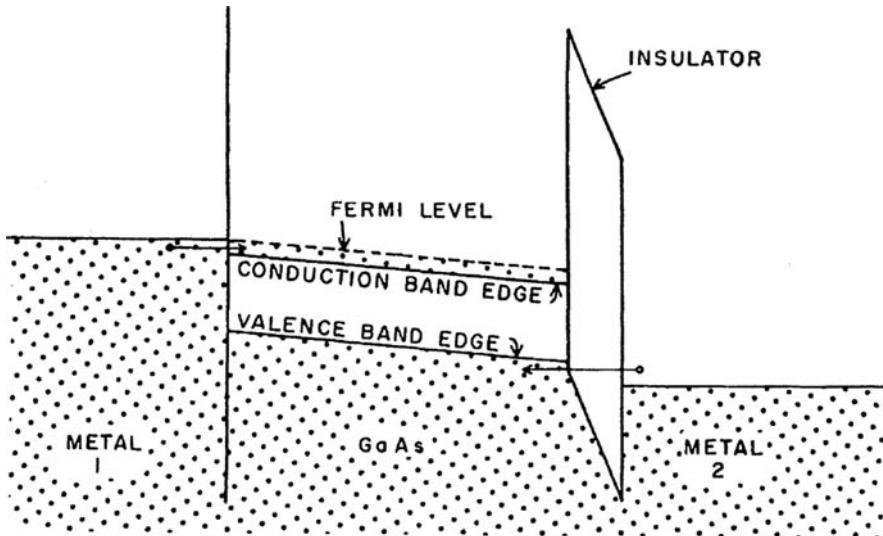


Fig. 12.7 Energy band diagram of a forward biased tunnel-injection laser

high electric field exists in this region. From the diagram, it can be seen that a hole in metal No. 2 does not have enough energy(classically speaking) to surmount the insulator barrier. However, quantum mechanical theory predicts that there is a small probability that the hole can pass through the insulator into the semiconductor even if classically it does not have enough energy to do so. In order for a significant number of holes to tunnel through, the insulator must be very thin or the voltage must be very high because the tunneling current density is given by [14, 15]

$$J_t = J_0 \left(\frac{E}{E_0} \right)^2 e^{-E_0/E}, \quad (12.23)$$

where E is the electric field in the insulator and J_0 and E_0 are constants, given by

$$E_0 = \frac{4\varphi^{3/2} (2m^*)^{1/2}}{3hq}, \quad (12.24)$$

and

$$J_0 = \frac{2q\varphi^2 m^*}{9h\pi^2}, \quad (12.25)$$

φ being the barrier height, and the other quantities are as defined previously. To obtain the necessary amount of current for laser threshold (about 100 A/cm^2), calculations show that a field strength of about 10^7 V/cm is needed. Thus, when using a 10 V source, one needs an insulator thickness of 10^{-6} cm or $0.01 \mu\text{m}$.

12.2.2 Lasing Threshold Conditions

The derivation of a threshold current relation for the tunnel-injection laser follows exactly the steps of (12.4), (12.5), (12.6), (12.7), (12.8), (12.9) and (12.10), except that the thickness of the light emitting layer D in the tunnel laser is equal to the thickness of the inverted population region d in the tunnel laser is equal to the thickness of the inverted population region D because the optical fields (or photons) are confined by reflection from the metal layers above and below the active layers. Thus, the value of D in (12.10) would be reduced by about a factor of 10 for the case of a GaAs tunnel laser as compared to that of a diode laser. In addition, the loss coefficient α is smaller in the case of the tunnel laser because interband absorption is suppressed in the active region by the inverted population.

As a numerical example, let us compare a conventional GaAs diode laser with $D = 10 \mu\text{m}$, $d = 1 \mu\text{m}$, and $\alpha = 35 \text{ cm}^{-1}$ to a GaAs tunnel laser with $d = D = 1 \mu\text{m}$ and $\alpha = 3 \text{ cm}^{-1}$. For both lasers, we take $R = 0.34$, $L = 1 \text{ mm}$, $\eta_q = 0.7$, $\lambda_0 = 9 \times 10^{-5} \text{ cm}$, $n = 3.34$ and $\Delta\nu = 1.5 \times 10^{13} \text{ s}^{-1}$. Then, from (12.10), we calculate

$$\begin{aligned} J_{\text{th}}(\text{diode}) &= 5.43 \times 10^3 \text{ A/cm}^2, \\ J_{\text{th}}(\text{tunnel}) &= 1.64 \times 10^2 \text{ A/cm}^2, \end{aligned}$$

and from (12.18) we find

$$\begin{aligned} \eta(\text{diode}) &= 24\%, \\ \eta(\text{tunnel}) &= 78\%. \end{aligned}$$

Because a very thin oxide layer is required in order to obtain sufficient tunneling current to exceed threshold, the tunnel-injection laser is relatively difficult to fabricate, and the heterojunction laser diode (to be discussed at length in the next chapter) is used in most applications. However, tunnel-injection has been shown to be particularly useful in oxide-confined vertical-cavity surface-emitting lasers (VCSELs) [16, 17]. As semiconductor fabrication technology has improved, work has continued on tunnel injection lasers grown in AlGaAs [18], GaInAsSb [19, 20], InGaAlAs [21], and GaInAs [22]. Tunnel-injection quantum dot lasers have been made [23]. (For a discussion of quantum dot structures, see Chapter 22.)

12.3 Polymer Lasers

The topic of polymer lasers already has been discussed in Chap.5, but it is worthwhile to reconsider the details of the lasing mechanism in a polymer now that we have seen its counterpart in a semiconductor, the electronic transition that results in stimulated emission and lasing is form a state near the bottom of the conduction

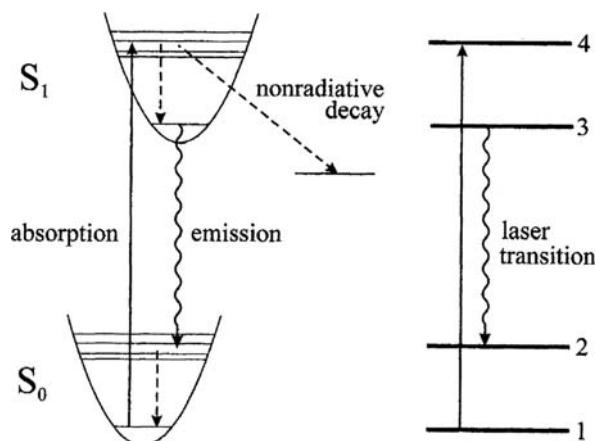


Fig.12.8 Energy band diagram[24]

band to a state near the top of the valence band. It has been found that polymers can have an effective optical band structure which is very similar to that of a semiconductor. For example, Schulzgen [24] has studied the polymer material BEH:PPV and has made lasers in it. The optical band structure of this material is shown in Fig. 12.8.

Optical pumping causes electrons to make upward transitions from S_0 states at energy level 1 to S_1 states at level 4. They then decay via a direct nonradiative transition to level 3. There also is some loss of photons via an indirect nonradiative transition. The lasing transition is between level 3 and level 2. Level 2 is depopulated via a direct nonradiative transition that returns electrons to the ground level 1. The peak wavelength of laser emission is 630 nm and the threshold excitation level is 40 $\mu\text{J}/\text{cm}^2$, as was seen in Fig.5.10.

12.4 New Semiconductor Materials for New Wavelengths

All of the semiconductor lasers that have been described so far have emitted the near-infrared wavelengths from about 800 to 1600 nm. However, in recent years, improvements in fabrication technology have made it possible to expand the range of wavelengths available from semiconductor lasers to both shorter and longer wavelengths.

12.4.1 Gallium Nitride Lasers

Blue-light lasers were long desired by engineers for data-recording applications, in which the shorter wavelength permits greater data densities, and for multicolor

displays. Because of its wide bandgap (3.5 eV) GaN was a logical choice for making a blue-light laser. The first blue-emitting (405-nm) laser was developed by Nakamura et al. in 1996 [25, 26]. In earlier attempts at making a GaN laser, epitaxial layers of GaN were grown on a sapphire substrate. That technique produced GaN crystalline layers that were oriented along the hexagonal c-plane. Unfortunately, strong polarization fields and piezoelectric effects occur along this plane that act to separate holes and electrons and prevent efficient recombination. Nakamura et al. began with a GaN substrate and grew the epitaxial layers along the rectangular m-plane, which has much lower polarization fields and piezoelectric effects. This “non-polar” GaN laser diode structure initially resulted in lasers that had threshold currents of approximately 7.5 kA/cm^2 when operated in pulsed mode at room temperature. Currently, non-polar GaN lasers emitting at 459 nm have been reported to operate in the cw mode at room temperature with a threshold current density of 5.0 kA/cm^2 and a threshold current of 40 mA [27].

12.4.2 Silicon Lasers

Because of its indirect bandgap silicon has not been thought of as a material for making efficient lasers. However, the well-developed commercial market for Si-based integrated circuits has fostered the dream of a Si-based OIC or PIC. Thus there have been efforts to make lasers in Si. Xu and Cloutier [28, 29] have reported that lasing at $1.28 \mu\text{m}$ was observed in a crystalline silicon structure under direct optical pumping. The lasing in Si was accomplished by incorporating a nanostructure into the light emitting region. (Nanostructures are discussed in Chapter 22.) The creation of a laser in Si is a promising development. However, optical pumping is not practical in many applications, and the fabrication of this Si laser involved sophisticated nanostructure self assembly growth techniques that are not yet suitable for production lines.

Engineers also are exploring the possibility of using the Raman effect to produce a laser in silicon. Rong et al. [30] have reported cw operation of a silicon Raman laser. In a Raman laser, a pumping source of light increases the vibrational energy of the atoms of a material. The interaction of the photons with these atoms causes the photons to lose energy, resulting in the generation of a secondary flux of photons of a longer wavelength. By proper choice of laser mirrors and pumping wavelength, a coherent beam of the secondary photons can be produced. In this case, the laser was constructed from a low-loss silicon-on-insulator (SOI) rib waveguide chip, the facets of which were coated with multilayer dielectric film mirrors. A p-i-n diode was imbedded in the waveguide and was reverse biased to reduce two photon absorption (TPA)-induced free carrier absorption (FCA). TPA-induced FCA reduces the gain of a Raman laser. The Raman laser wavelength was 1,686 nm and the pump wavelength was 1,550 nm. The slope efficiency for single side output was 4.3% for 25 V bias and 3% for 5 V bias. At 25 V bias the curve of output power vs. input power was linear up to 7 mW output.

Problems

- 12.1 We would like to design a p-n junction laser for use as the transmitter in a range finder. The output is to be pulsed with a peak power out of each end of 10 W (only the output from one end will be used) and a pulse duration of 100 ns. Wavelength is to be 9000 Å. Room temperature operation is desired and some of the pertinent parameters have been either measured or established so that:
- (1) Half-power points of the emission peak for spontaneous emission have been measured for this material at room temperature to be at 9200 and 8800 Å.
 - (2) Index of refraction: 3.3.
 - (3) Thickness of light inverting layer: 10 μm.
 - (4) Thickness of active (inverted pop.) layer: 1 μm.
 - (5) Internal quantum efficiency: 0.7.
 - (6) Average absorption coefficient: 30 cm^{-1} .
 - (7) $W = 300 \mu\text{m}$.
 - (8) Reflectivity of the Fabry-Perot surfaces: 0.4.
 - (a) What must be the separation between Fabry-Perot surfaces if we wish to have a peak pulse current density of $3 \times 10^4 \text{ A/cm}^2$?
 - (b) What is the threshold current density?
- 12.2 In the case of the laser of Problem 12.1, if the heat sink can dissipate 1 W at room temperature, what is the maximum pulse repetition rate that can be used without causing the laser crystal to heat above room temperature (neglect I^2R loss)?
- 12.3 For the laser of Problem 12.2, what is the minimum range for which transmitted and target reflected pulses will not overlap in time at the detector, assuming that transmitter and detector are located at essentially the same point and transmission is through air?
- 12.4 (a) Can a light emitter without mirrors or any other optical feedback device produce light by stimulated emission? Is this light coherent?
(b) Explain the significance of the *threshold* phenomenon in lasers.
(c) Why do confined-field type lasers have lower threshold current and higher efficiency?
(d) Why is the divergence angle of the emitted optical beam much greater in a semiconductor laser than it is in a gas laser?
- 12.5 A semiconductor laser formed in a direct bandgap material is found to have an emission wavelength of 1.2 μm. The external quantum efficiency is 15%.
- (a) What is the approximate bandgap energy of the material?
 - (b) If the output power is 20 mW, give an approximate estimate of the input current.

- 12.6 A confined-field laser has been properly designed so that its absorption loss due to interband transitions is negligible ($\alpha_{IB} \cong 0$). The free carrier absorption loss and the scattering loss are $\alpha_{FC} = 5 \text{ cm}^{-1}$ and $\alpha_S = 0.5 \text{ cm}^{-1}$, respectively. If the length of the laser is increased by a factor of two, by what factor will the threshold current density change? (Assume the end face reflectivity is 65%).
- 12.7 Derive an expression for the threshold current density J_{th} of an injection laser which has one end face coated with a layer that is totally reflecting ($R = 1$). Assume that the relationship between the input current density $J[\text{A}/\text{cm}^2]$ and the gain coefficient $g[\text{cm}^{-1}]$ is given by $J = Kg$, where K is a constant. Also assume that R_0 is the reflectivity of the partially transmitting endface, and that L , α , D , and d are as previously defined.
- 12.8 The following parameters are known for a semiconductor laser:
- Emission wavelength $\lambda_0 = 0.850 \mu\text{m}$.
 - Lasing threshold current $I_{th} = 12 \text{ A}$
 - External quantum efficiency = 1% below threshold*
 - External quantum efficiency = 10% above threshold*
- (a) Sketch the light output power [W] vs. input current [A] curve for this device for currents ranging from 0 to 20 A; label the value of the output power at threshold.
- (b) What is the output power [W] for an input current = 18 A?
- 12.9 A multimode GaAs laser diode is emitting at a wavelength of approximately $0.9 \mu\text{m}$. At that wavelength the index of refraction of GaAs is 3.6 and the dispersion is $0.5 \mu\text{m}^{-1}$. The laser cavity length of $300 \mu\text{m}$. Calculate the mode spacing (in wavelength) at approximately $0.9 \mu\text{m}$.
- 12.10 The external *differential quantum efficiency* of a GaAs laser diode emitting at a wavelength of $0.9 \mu\text{m}$ is 30%. The applied voltage is 2.5 volts. Calculate the external *power efficiency* of the device.
- 12.11 A certain semiconductor p-n junction laser when biased as required has a gain coefficient $g = 70 \text{ cm}^{-1}$, a loss coefficient $\alpha = 5 \text{ cm}^{-1}$, and a length between Fabry-Perot end faces of $L = 300 \mu\text{m}$. It is desired to add anti-reflection coatings to the endfaces. What is the smallest endface reflectivity R that can be used without going below the lasing threshold?
 (Assume that both endfaces have the same reflectivity, and that the confinement factor $d/D = 1$.)
- 12.12 A p-n junction, Fabry-Perot type laser of cavity length L and mirror reflectance R is to be redesigned to double L , but to maintain the same threshold current density. All other junction and material parameters remain the same.
- (a) What is the new mirror reflectance?
- (b) What is the new threshold current in terms of the old threshold current (ie. a ratio)?

- 12.13 In a particular tunnel injection laser an electric field in the insulator of 4×10^4 V/m produces a current density of 1500 A/cm^2 . If the thickness of the insulator is reduced by a factor of $\frac{1}{2}$, by what factor will the current density be increased?

References

1. B.N. Hall, G.E. Fenner, J.D. Kingsley, T.J. Soltys, R.O. Carlson: Phys. Rev. Lett. **9**, 366 (1962)
2. M.I. Nathan, W.P. Dumke, G. Burns, F.H. Dill Jr., G. Lasher: Appl. Phys. Lett. **1**, 62 (1962)
3. T.M. Quist, R.H. Rediker, R.J. Keyes, W.E. Krag, B. Lax, A.L. McWhorter, H.J. Zeiger: Appl. Phys. Lett. **1**, 91 (1962)
4. N. Holonyak Jr. S.F. Bevacqua: Appl. Phys. Lett. **1**, 82 (1962)
5. A. Yariv: *Optical Electronics*, 4th edn. (Holt, Rinehart and Winston, New York 1991) p. 116
6. G. Wade, C.A. Wheeler, R.G. Hunsperger, T.O. Carroll: A tunnel injection laser. 5th Int'l Cong. on Microwave Tubes, Paris, France (1964)
7. J. Lasher: IBM J. **7**, 58 (1963)
8. G. Wade, C.A. Wheeler, R.G. Hunsperger: IEEE Proc. **53**, 98 (1965)
9. G. Diemer, B. Bölgger: Physica **29**, 600 (1963)
10. T. Pecany: Phys. Stat. Sol. **6**, 651 (1964)
11. H. Kroemer: IEEE Proc. **51**, 1782 (1963)
12. Zh.I. Alferov: Sov. Phys.-Solid State **7**, 1919 (1966)
13. H. Kressel, N. Nelson: RCA Rev. **30**, 106 (1969)
14. R.H. Fowler, L. Nordheim: Proc. Roy. Soc. (London) A **119**, 173 (1928)
15. A.G. Chynoweth: Progr. Semiconduct. **4**, 97 (1960)
16. D.L. Huffaker, D.G. Deppe: Improved performance of oxide-confined vertical-cavity surface-emitting lasers using a tunnel injection active region, Appl. Phys. Lett. **71**, 1449 (1997)
17. D.L. Huffaker, D.G. Deppe: Intractivity Contacts for Low-Threshold Oxide-Confined Vertical-Cavity Surface-Emitting Lasers, IEEE Phot. Tech. Lett. **11**, 934 (1999)
18. B.A. Vojak, N. Holonyak, Jr., R. Chin, E.A. Rezek, R.D. Dupuis, P.D. Dapkus: Tunnel injection and phonon assisted recombination in multiple quantum-well AlGaAs-GaAs p-n heterostructure lasers grown by metalorganic chemical vapor deposition, Appl. Phys. **50**, 5835 (1979)
19. Yu.P. Yakovlev, K.D. Moiseev, M.P. Mikhailova, O.G. Ershov, G.G. Zegrya: Advanced tunnel-injection laser based on the type II broken-gap GaInAsSb/InAs heterojunction for the spectral range 3–3.5 μm , Digest of Conference on Lasers and Electro-Optics, CLEO '96. (Anaheim, CA, June 2–7, 1996)
20. K.D. Moiseev, M.P. Mikhailova, O.G. Ershov, Yu. P. Yakovlev: Tunnel-injection laser based on a single p-GaInAsSb/p-InAs type-II broken-gap heterojunction, Semiconductors, **30**, 223 (1996)
21. P. Podemski, R. Kudrawiec, J. Misiewicz, A. Somers, J.P. Reithmaier, A. Forchel: On the tunnel injection of excitons and free carriers from $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{In}_{0.53}\text{Ga}_{0.23}\text{Al}_{0.24}\text{As}$ quantum well to $\text{InAs}/\text{In}_{0.53}\text{Ga}_{0.23}\text{Al}_{0.24}\text{As}$ quantum dashes, Appl. Phys. Lett. **89**, 061902 (2006)
22. M. Ohta, T. Furuhata, T. Iwasaki, T. Matsuura, Y. Kashihara, T. Miyamoto, F. Koyama: Structure-dependent lasing characteristics of tunnel injection GaInAs/AlGaAs single-quantum-well lasers, Jpn. J. Appl. Phys. **45**, L162 (2006)
23. Yu.M. Shernyakov, A.Yu. Egorov, A.E. Zhukov, S.V. Zaitsev, A.R. Kovsh, I.L. Krestnikov, A.V. Lunev, N.N. Ledentsov, M.V. Maksimov, A.V. Sakharov, V.M. Ustinov, C. Chen, P.S. Kop'ev, Zh.I. Alferov, D. Bimberg: Quantum-dot cw heterojunction injection laser operating at room temperature with an output power of 1 W, Tech. Phys. Lett. **23**, 149 (1997)

24. A. Schulzgen, C. Spiegelberg, S.B. Mendes, P.M. Allemand, Y. Kawabe, M. Kuwata-Gonokami, S. Honkanen, M. Fallahi, B.N. Kippelen, N. Peyghambarian: Light amplification and laser emission in conjugated polymers. Opt. Eng. **37**, 1149 (1998)
25. S. Nakamura, G. Fasol: *The Blue Laser Diode – GaN based Light Emitters and Lasers* (Springer, Heidelberg, 1997)
26. S. Nakamura, S. Pearton, G. Fasol: *The Blue Laser Diode – The Complete Story* 2nd updated and extended ed. (Springer, Heidelberg, 2000)
27. M. Kubota, K. Okamoto, T. Tanaka, H. Ohta: Continuous-wave operation of blue laser diodes based on nonpolar m-plane gallium nitride, Appl. Phy. Express **1**, 011102 (2008)
28. J. Xu: Directly pumped crystalline silicon laser – an impossible possibility?, Digest 3rd IEEE International Conference on Group IV Photonics 2006, 213 (2006)
29. S.G. Cloutier, J.M. Xu: All-silicon laser, Condensed Matter Archives, Arxiv cond-mat 0412376 (2004)
30. H. Rong, R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang, M. Paniccia: A continuous-wave Raman silicon laser, Nature **433**, 725 (2005)

Supplementary Reading on Semiconductor-Laser Fundamentals

- N.W. Carlson: *Monolithic Diode-Laser Arrays*, Springer Ser. Electron. Photon., Vol. 33 (Springer, Berlin, Heidelberg 1994)
- W.W. Chow, S.W. Koch, M. Sargent III: *Semiconductor-Laser Physics* (Springer, Berlin, Heidelberg 1994)
- C.F. Klingshirn: *Semiconductor Optics* (Springer, Berlin, Heidelberg 1995)
- H. Kressel, M. Ettenberg, J.P. Wittke, I. Ladany: Laser diodes and LEDs for fiber optical communications, in *Semiconductor Devices for Optical Communications*, H. Kressel 2nd edn., (ed.) (Springer, Berlin, Heidelberg 1982) pp. 9–62
- J. Pankove: *Optical Processes in Semiconductors* (Prentice-Hall, Reading, MA 1971) Chap. 10
- A. Yariv: **Quantum Electronics**, 3rd edn. (Wiley, New York 1989) pp. 232–263
- P. Yu, M. Cardona: *Fundamentals of Semiconductors: Physics and Material Properties* (Springer, Berlin, Heidelberg 1995)

Chapter 13

Optical Amplifiers

In lightwave communications systems, just as in wired communication systems, there is a need for repeaters at regular intervals to amplify the signal to compensate for losses. In early lightwave systems this function was accomplished by using a photodetector to convert the lightwave signal to an electrical current waveform, amplifying it electronically, and then converting it back to an optical form with a laser or high-speed LED. This approach involved the use of additional components, which inevitably reduced the overall reliability of the system since each component has a certain probability of premature failure. Electronic amplification also limited the overall bandwidth of the lightwave system. In order to overcome these problems, researchers have studied and developed a number of different types of optical amplifiers, devices that directly amplify the optical signal without converting it to electronic form. The most commonly used optical amplifiers are the erbium doped fiber amplifier (EDFA), the semiconductor optical amplifier (SOA) and the Raman fiber amplifier (RFA). These devices are discussed in the following sections of this chapter.

13.1 Optical Fiber Amplifiers

It has been found that the glass optical fibers that are used as waveguides in communications and datacom systems can also be used as optical amplifiers if they are doped with optically active atoms such as Erbium. These atoms enter the structure of the host glass as ions that have a characteristic splitting of ion energy levels that is suitable for producing stimulated emission of photons when the system is optically pumped at an appropriate wavelength and is stimulated by a lightwave signal. The number of photons produced by stimulated emission is proportional to the number of photons present in the stimulating lightwave signal. Thus, the amplitude modulation(information signal) of the amplified lightwaves is preserved. Because of their relative efficiency and reliability, optical fiber amplifiers have become the predominant type of amplifier used in long-distance lightwave telecom and datacom systems.

13.1.1 Erbium Doped Fiber Amplifiers

The erbium doped fiber amplifier(EDFA) is the one that is most widely used because the energy levels of the Er^{3+} ion in a glass host provide a convenient pumping wavelength centered about $0.94 \mu\text{m}$ and a stimulated emission central wavelength of $1.54 \mu\text{m}$, corresponding to the third “telecommunications windows”. Glass fibers have three wavelength regions of low loss at approximately $0.8 \mu\text{m}$, $1.3 \mu\text{m}$ and $1.55 \mu\text{m}$ that have come to be known as the first, second and third telecommunications windows, respectively. Semiconductor lasers, which are commercially available with emission wavelengths of 810 and $980 \mu\text{m}$, are generally used as the pumping source. However, there is also a strong absorption band centered about $1.5 \mu\text{m}$, which permits pumping with $1.48 \mu\text{m}$ semiconductor lasers. The absorption characteristics of the Er^{3+} ion in a silica host are shown in Fig.13.1.

The transitions that are involved in producing the stimulated emission of an EDFA are shown in Fig.13.2. Most often a pump wavelength of 980 nm is used to raise ions to the upper level. The lifetime in this state is very short, on the order of $1 \mu\text{s}$. Ions then decay to a metastable state with a relatively long lifetime of approximately 10 ms . This long lifetime allows the population of ions in the metastable state to build up so as to produce an inverted population with respect to the ground state. Passing photons of the lightwave signal can then reproduce themselves by stimulated emission of a duplicate photon, as an ion makes the transition to the ground state are not single energy levels, but rather are manifolds or bands of energy levels, is very important to the operation of an EDFA. It results in the amplifier having gain and therefore being useful over a range of wavelengths from approximately 1520 to 1600 nm .

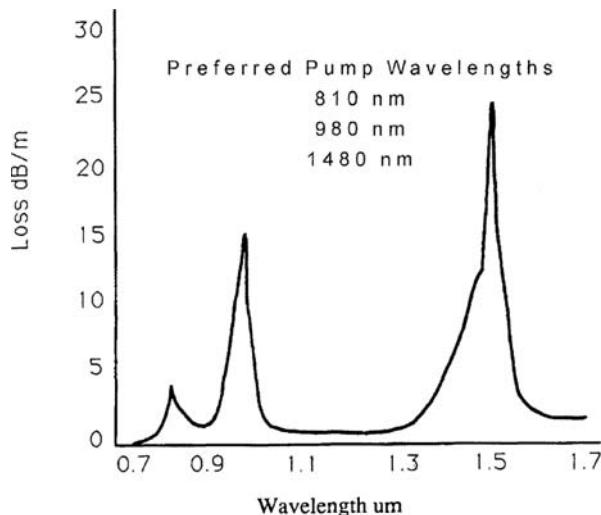


Fig.13.1 Absorption spectrum of erbium-doped silica fiber Reprinted from [1] p.146 by courtesy of Marcel Dekker, Inc.

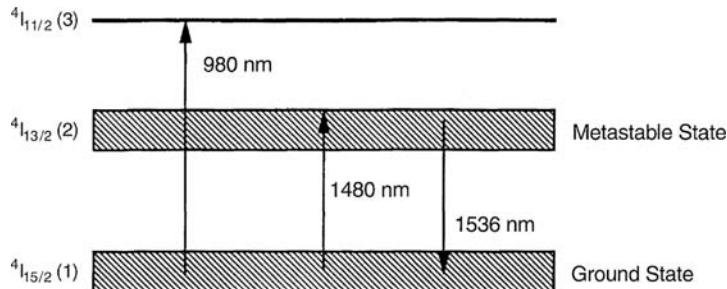


Fig.13.2 Energy levels and transitions in an EDFA Reprinted from [1] p.147 by courtesy of Marcel Dekker, Inc

The physical implementation of an EDFA is particularly convenient in that, being a fiber, it can be directly inserted into a lightwave communication line with minimal coupling loss. The pumping photons are usually supplied by a semiconductor diode laser as shown in Fig.13.3.

The gain of an optical fiber amplifier is proportional to the degree of population inversion and can be represented by:

$$G_{\text{opt}} = C(N_2 - N_1), \quad (13.1)$$

where N_2 is the number of ions in the upper energy level of the gain-producing transition, N_1 is the number in the lower level and C is a constant that represents the probability of a transition. The quantity CN_1 accounts for the possibility of upward, or absorptive, transitions. In order to achieve overall gain, the system must be pumped to make N_2 significantly greater than N_1 . The optical power (photons/s) generated by stimulated emission is given by:

$$P_{\text{stim}} = P'_{21} N_2 I_p, \quad (13.2)$$

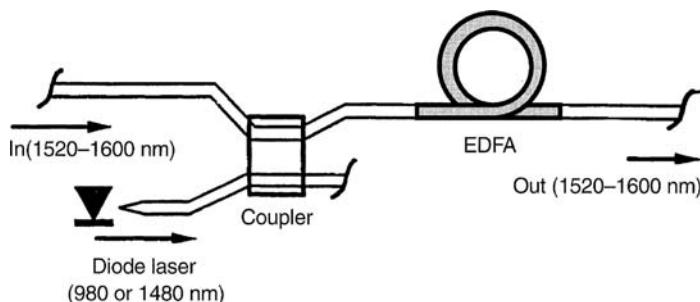


Fig.13.3 Physical configuration of an EDFA

where P'_{21} is the probability of a downward stimulated transition and I_p is the photon intensity, or density of photons (photons/s) with the “right” energy (that corresponding to the difference between upper and lower energy levels). Unfortunately, we also have to consider power generated due to amplified spontaneous emission (ASE), which is noise and is not part of the desired amplified signal

$$P_{ase} = P_{21}N_2. \quad (13.3)$$

where P_{21} is the probability of a downward spontaneous transition. In general it would also be necessary to consider power lost due to absorption P_{abs} which is given by:

$$P_{abs} = P_{12}N_1, \quad (13.4)$$

where P_{12} is the probability of an upward, absorptive transition. Therefore the optical gain would be expressed as:

$$G_{opt} = (P_{out} - P_{ase})/P_{sig} = (P_{stim} - P_{abs} - P_{ase})/P_{sig}, \quad (13.5)$$

where P_{sig} is the input optical signal to be amplified. However, for a strongly inverted population $N_2 \gg N_1$, and absorption can be neglected. Thus, the expression simplifies to:

$$G_{opt} = (P_{stim} - P_{ase})/P_{sig}, \quad (13.6)$$

The gain is frequently expressed in decibels (dB),

$$G_{opt} = 10 \log [(P_{stim} - P_{ase})]/P_{sig}. \quad (13.7)$$

A gain of 30 db (10^3) is easily attainable in an EDFA with only a few mW of pump power, P_{pump} , at 980 nm. Pumping at 980 nm is more efficient than pumping at 1480 nm. A gain efficiency (G_{opt}/P_{pump}) of 8–10 dB/mW is typical for pumping at 980 nm.

EDFAs are very stable with respect to ambient environmental conditions. Liu et al. [2] have reported that the temperature coefficient of EDFA gain is 0.023 dB/ $^{\circ}$ C in the temperature range from 25 to 95 $^{\circ}$ C for pumping at 980 nm. Yamada et al. [3] found that the temperature coefficient (neglecting loss) for similar pumping conditions was less than 0.013 dB/ $^{\circ}$ C for a fiber shorter than 12 m in the range from –40 to +50 $^{\circ}$ C. The effect of electric field, which could change the Stark Effect splittings of the energy band manifold of the Er³⁺ ion, were found to be negligible for electric field strength less than 3 kV/m[2]. The effect of magnetic fields was also reported to be small. A magnitude of 200 G only slightly increased gain [2]. The robustness and efficiency of the EDFA at wavelengths within the third telecommunications window have made it the most widely used type of optical fiber amplifier

(OFA). However, other OFAs have been developed in order to expand the range of wavelength available for dense wavelength division multiplexing (DWDM).

13.1.2 Raman Optical Fiber Amplifiers

The range of wavelengths over which an EDFA can be used is relatively limited. It can be used with maximum efficiency only over the range from approximately 1525–1575 nm, within which glass optical fiber loss is minimal. The range can be extended to longer wavelengths out to 1600 nm, with reduced gain and increased loss, but there is no significant gain at shorter wavelengths. This has led researchers to develop new types of OFAs to take advantage of the shorter optical wavelengths that are still within the third telecommunications window. The absorption characteristics of currently available glass optical fibers in this wavelength range have been summarized by Islam and Nietubyc [4], as shown in Fig. 13.4.

Attenuation curves are shown for standard single-mode fiber (SSMF) and for Lucent Allwave, an OH⁻ content-reduced, single-mode fiber. Modern low OH⁻ content silica fibers have less than 0.35 dB/km loss over the wavelength range from 1300 to 1600 nm, permitting the distribution of DWDM channels over the entire breadth of the S, C and L optical communication bands, as long as suitable amplifiers can be found to make the required repeaters.

The Raman fiber amplifier (RFA) can extend the range of useful wavelengths into the S band. Conservation of energy is still required, as it was in the case of standard

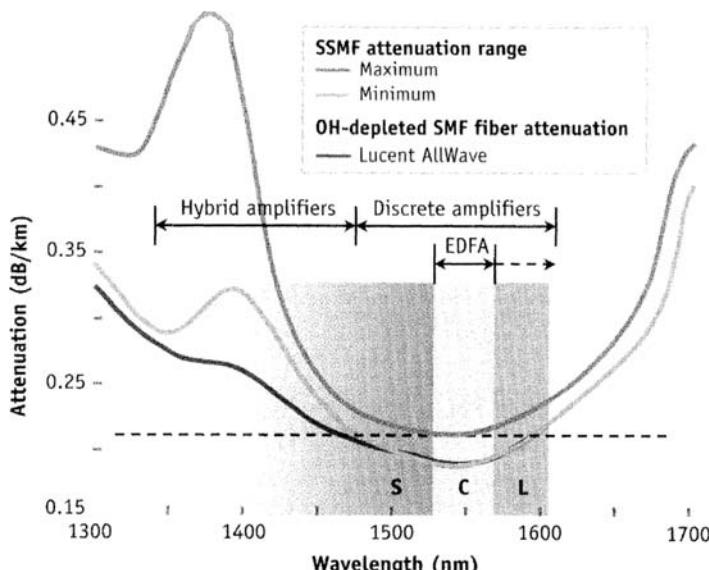


Fig. 13.4 Absorption characteristics of glass optical fiber waveguides [4]

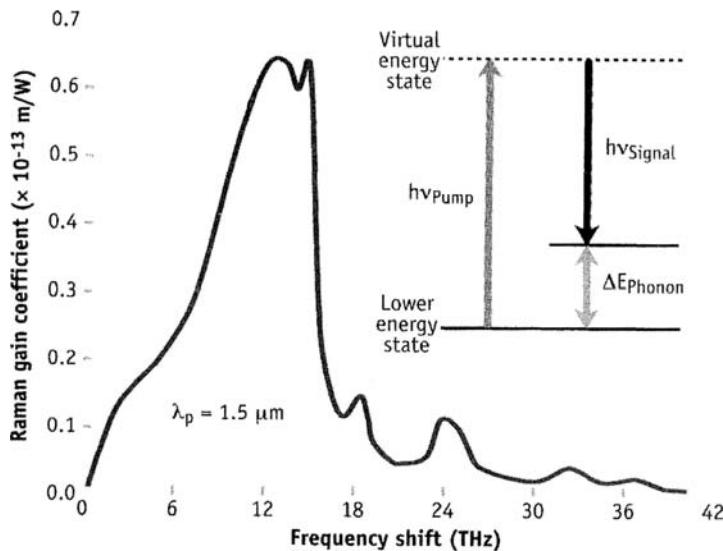


Fig. 13.5 Raman gain spectrum [4]

stimulated emission, but during Raman amplification additional energy is provided by phonons. Power is shifted into the signal optical beam, which is at a lower frequency than the pump beam. Photons of the pumping beam interact with phonons to downshift power into the optical signal beam, while energy and momentum are both conserved. When pumped with a laser source of 1500 nm wavelength, the Raman gain spectrum of a silica fiber is over 40 THz wide, as shown in Fig.13.5.

It has a dominant peak near + 13.2 THz (or approximately 100 nm offset). The output wavelength peaks in the S-band at 1480 nm. The RFA can be pumped with laser diodes, but both the pump and the signal beams must have the same polarization.

13.1.3 Other Optical Fiber Amplifiers

Another approach that has been used to produce amplifiers that are useful at shorter wavelengths is to dope the glass fiber with different atoms. For example, Thulium-doped fiber amplifiers (T DFA) have been used for the S-band (1480–1520 nm) [4] and also for the first telecommunications window near 1.3 μm . TDFAs have a three level energy system that requires two optical pump sources. An upconverting pump laser is coupled into the amplifier fiber in the forward direction to pump the Tm^{3+} ions from the ground state to an intermediate energy level (E_2). A second counterpropagating pump light source raises the ions to a higher energy level (E_3). Amplification then occurs for the signal photons, which have energy $E_{\text{signal}} = E_3 - E_2 = h\nu_{\text{signal}}$. The result is a gain response that extends from approximately

1430–1500 nm, with a peak at 1460 nm. Praseodymium-doped fiber amplifiers have a broad gain curve that is peaked at 1300 nm and extends from approximately 1260–1340 nm [1]. They exhibit strong gain at 1300 nm, but the lifetime of Pr^+ ions in a silica host is very short; hence, efficiency is low. To circumvent this problem, fluoride fibers such as ZBLAN are used. However, even with these a pump power of 400–600 mW at 1010 nm wavelength is required to achieve 30 dB of gain. By comparison, an EDFA needs only 3 mW. Praseodymium lasers can be pumped at 1010 nm by Nd:YLF lasers.

13.2 Non-Fiber Ion-Doped Optical Amplifiers

The use of ion-doping to produce light emitters and even lasers in polymer materials was discussed in Chapter 5. (See, for example, the input and output curves of a doped polymer laser in Fig. 5.10.) It is therefore not surprising that ion-doping can be used to produce optical amplifiers in waveguiding layers of dielectric materials, as well as glass fibers. Such planar optical amplifiers, which are particularly useful in optical integrated circuits, have been produced by a number of different researchers [5–11]. For example, Jiang et al. [5] have produced an Er^{3+} doped amplifier in a new rare-earth phosphate glass designated as type MM-1. The waveguides were of the ion-exchange type, made by immersing the glass substrate in $\text{KNO}_3 + \text{AgNO}_3$ for 120 minutes at 365°C. The resulting waveguiding layers were single-mode at a wavelength of 1.54 μm , and three-mode at 632.8 nm. The Er^{3+} ion in this glass host material has a broad absorption peak at 1480 nm, which can be used for pumping, and an emission (gain) curve extending from about 1500–1575 nm, with a peak at 1530 nm.

13.3 Semiconductor Optical Amplifiers

The double heterostructure, p-n junction diode which is used in semiconductor lasers can also function as an optical amplifier. Operating in this mode, the device increases the optical power of light entering one of its facets to produce a higher power output at the other facet. Intensity modulation variations are preserved and amplified, so the diode can be used as a repeater in optical communication systems. The basic mechanism of amplification is that of stimulated emission, just as in a laser. However, in the amplifier the diode is usually biased somewhat below lasing threshold so that oscillation does not occur.

There are three types of semiconductor laser amplifiers: the traveling-wave type, the Fabry-Perot type, and the injection-locking type. In the traveling-wave type of laser amplifier the facets are antireflection coated so that the light waves make only one pass through the inverted population region of the diode. In GaAs and its related ternary and quaternary compounds, the reflectivity of a cleaved endface is approximately 33%. However, it can be reduced to only a percent or two by the application

of a properly designed multilayer dielectric antireflection coating. In the Fabry-Perot type of laser amplifier the end-facets are left uncoated. Thus incident light is amplified during successive passes between the end-facets before being emitted at a higher power level. The third type of laser amplifier, the injection-locked amplifier, is different from the other two types in that it is biased above threshold and actually operates as a laser. However, the injection of an intensity modulated light signal at one facet causes the light waves produced by the lasing action to "lock" with the input light waves so as to follow the same intensity and phase variations, but at a higher optical power level. In all three types of laser amplifier the source of the added power is, of course, the electrical energy introduced by the bias power supply. The advantages of the semiconductor optical amplifier are that it is fast-switching, can cover the entire wavelength range from 800 to 1600 nm, is small, light and can be conveniently packaged as shown in Fig. 13.6.

The single-pass gain of the Fabry-Perot laser amplifier and the traveling-wave amplifier are given by [12]

$$G_s = \exp [(\Gamma g(N) - \alpha) L], \quad (13.8)$$

where Γ is the confinement factor, α is the internal loss coefficient, and L is the length of the cavity. The gain function, $g(N)$ depends linearly on the carrier concentration N so that

$$g(N) = aN - b, \quad (13.9)$$

where a and b are constants.

The injection current is given by

$$I = NeV/\tau_s \eta_i, \quad (13.10)$$

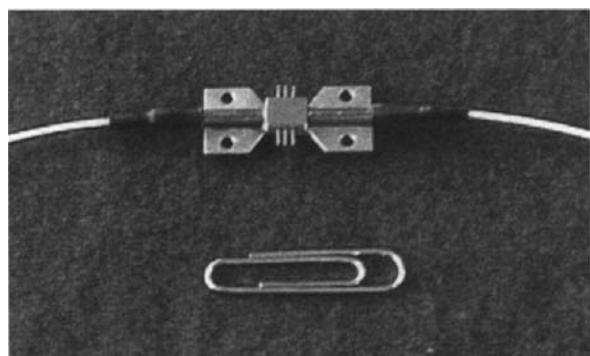


Fig. 13.6. Packaged semiconductor optical amplifier

where e is the electronic charge, V is the active volume (not the voltage), τ_s is the carrier lifetime and η_i is the internal quantum efficiency.

For an ideal traveling wave laser amplifier the total gain is just the gain per single pass G , as given in (13.8). For a Fabry-Perot laser amplifier the gain is enhanced by multiple reflections so that total gain is given by [13].

$$G = (1 - R_1)(1 - R_2) G_s / \left(1 - (R_1 R_2)^{1/2} G_s\right)^2. \quad (13.11)$$

where R_1 and R_2 are the endface reflectivities. It might appear that the Fabry-Perot laser amplifier would offer higher gain than the traveling-wave type. However, it must be remembered that the maximum value of G_s , (corresponding to the point of lasing threshold) is much larger for a traveling-wave laser amplifier because of the greatly reduced endface reflectivities. The gain of both the traveling-wave and Fabry-Perot laser amplifiers is limited by the fact that they cannot be operated above threshold. However, because of the reduced reflectivity of the mirrors the threshold level is much higher for the traveling-wave device.

Noise also limits the useful gain of these amplifiers, and that of the injection-locked laser amplifier as well. The major source of noise is the background of photons spontaneously emitted in the p-n junction [13]. This light is amplified along with the desired signal, reducing the signal-to-noise ratio. In the Fabry-Perot laser amplifier the cavity resonances reinforce and amplify only those frequencies of light which correspond to an integral number of half wavelengths between the two endfaces. The emission spectrum of the amplifier therefore contains a central longitudinal mode, the amplified signal, along with side lobes spaced on the order of 1 nanometer apart corresponding to other longitudinal mode resonances within the cavity. When the amplified output is later detected by a photodiode the various modes combine to generate what is known as “beat” noise. This noise adds to the amplified spontaneous emission noise mentioned previously. The noise spectrum generally spreads over a wider range of wavelengths than does the amplified signal. Thus the signal-to-noise ratio of these devices can be improved by inserting a narrow-bandpass filter at the output.

The Fabry-Perot laser amplifier is suitable for only low gain applications because its maximum gain is limited by the onset of laser oscillation at a relatively low threshold. The frequency of the input signal must be precisely matched to that of the main Fabry-Perot mode for maximum gain and signal-to-noise ratio; thus its bandwidth is limited. The injection-locked laser amplifier also has a limited bandwidth because the frequency of the input lightwaves must be close to that of the laser oscillation frequency if locking is to occur. The traveling-wave laser amplifier has an advantage over the others in that it has high gain, wide spectral bandwidth and low sensitivity to changes in either temperature or drive current level. For example, Zhang et al. [14] have reported traveling-wave laser amplifiers in GaAlAs with a relatively flat gain of over 20 dB for wavelengths from 886.1 to 868.5 nm wavelength.

13.3.1 Integrated Semiconductor Optical Amplifiers

Semiconductor optical amplifiers can be integrated with other devices in an optical integrated circuit. For example, Brenner et al. [15] have made InGaAsP/InP traveling wave amplifiers with high gain and integrated them with butt-coupled waveguides for OIC's. Johnson et al. [16], working in InGaAsP, have monolithically integrated a semiconductor optical amplifier, electroabsorption modulator and dual-waveguide mode spotsize converter all on an InP substrate. Their device is shown in Fig. 13.7.

It is fabricated by using a five-step MOVPE process. Two of the steps are selective-area growth steps in which SiO_2 is used to mask the growth. A unique feature of the device is a dual-waveguide, mode spot size converter that is designed to provide efficient coupling with the waveguides of a silica planar lightguide circuit to which the InP chip is to be flip-chip bonded. The resulting OIC is a wavelength-selectable laser (WSL), consisting of a hybrid array of DFB lasers, each capable of operating at a different wavelength when selected, sharing a single electroabsorption modulator. Top and side views of the spot size converter are shown in Fig. 13.8.

At the lateral taper the mode transfers from the upper waveguide to the lower waveguide, which then tapers in thickness to expand the mode. When tested in a cleaved, single-mode fiber link, the device exhibited greater than 10 dB fiber-to-fiber gain, with less than 4 dB coupling loss and output power of +4 dBm in saturation. The EA modulator exhibited a modulation bandwidth of 6 GHz and extinction ratios of 20 dB dc and 14.4 dB rf for a 2.4 V_{pp} drive.

Semiconductor optical amplifiers are particularly useful in integrated array applications because of their small size and compatibility with other optoelectronic devices. The WSL described above is a good example of that type of application.

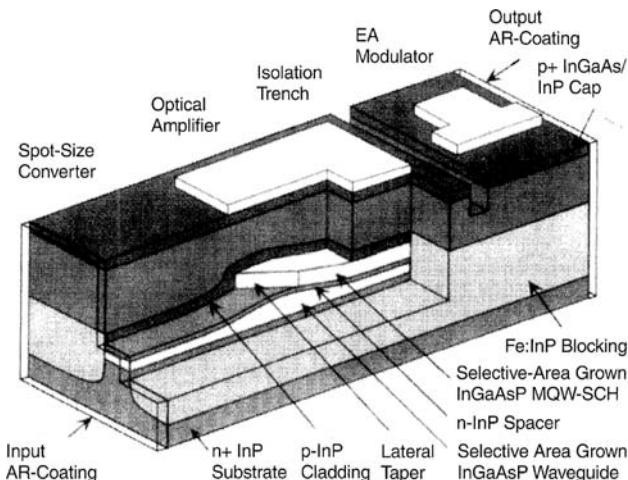


Fig. 13.7 Integrated SOA, EO modulator and mode shape converter [16] © 2000 IEEE

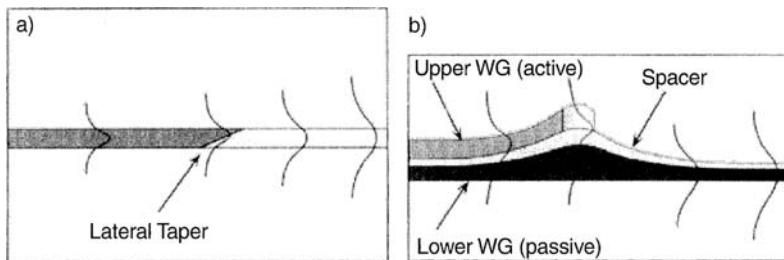


Fig. 13.8 a,b Spot size converter (a) top, (b) side [16] © 2000 IEEE

Another array application involving SOAs is the eight-channel bent waveguide SOA array described by Hatakeyama et al. [17]. Their array consisted of eight channels incorporating spot-size-converter and SOAs, with uniformly high performance. The fiber-to-fiber gain was 12.7 dB. When operated as a switching gate array, the channels had an on-off ratio greater than 50 dB (+10 dB on at 40 mA drive current to -40 dB when off at 0 mA).

13.4 Comparison of Ion-Doped Fiber Amplifiers with SOAs

13.4.1 Wavelength Range

Of primary importance in the comparison of different types of optical amplifiers is the wavelength range over which they can be used. To maximize the number of channels that can be used in a DWDM system it will be necessary to take advantage of all three of the “windows” of low loss in silica optical fibers. Semiconductor optical amplifiers can be used over the entire range from 800 nm to 1600 nm, while different types of ion-doped lasers must be chosen for different wavelength ranges. A summary of the types of amplifiers that are appropriate for different wavelength ranges is given in Table 13.1.

13.4.2 Performance Characteristics

The performance characteristics of SOAs can be easily compared to those of EDFA because both of these two amplifiers have been widely studied and used in a variety of applications. They are both commercially available. Other types of ion-doped amplifiers, such as thulium or neodymium-doped are reported in the literature but have not yet been well developed. A fairly thorough comparison of EDFA and SOAs has been done by Fake and Parker [1]. Their results are given in Table 13.2.

The characteristics shown for the EDFA probably also apply generally to the other types of ion-doped amplifiers. As these devices are developed further, a more accurate comparison will be possible.

Table 13.1 Optical communications bands and compatible amplifiers

First Window	(approximately 800–900 nm)
	SOA
	TDFA (Thulium)
Second Window	(approximately 1300–1360 nm)
	SOA
	PDFA (Praseodymium)
	NDFA (Neodymium)
Third Window	
S-band	(1480–1520 nm)
	SOA
	TDFA
	Raman
C-band	(1525–1565 nm)
	SOA
	EDFA
	Raman
L-band	(1570–1620 nm)
	SOA
	Raman

Table 13.2 Performance characteristics of EDFAs and SOAs Reprinted from [1] p.165 by courtesy of Marcel Dekker, Inc.

EDFA Character	Explanation	SOA Characteristics	Explanation
High gain (40 dB) and saturation output power	High pump power Efficient fiber Long interaction length	Low gain (15 dB) and saturation output power	Short cavity reflections Poor coupling efficiency
Good noise figure (3.5 dB)	Low splice loss to fiber High pump power Negligible internal reflections	Poor noise figure	Reflections (ripple) Coupling efficiency
Polarization insensitive	Circular fiber geometry	Polarization sensitive	Rectangular waveguide geometry favors TE mode
Cross-talk negligible (modulation > 10 kHz)	Lifetime long 10 mS	Cross-talk significant at useful data rates (< 10 Gb/s)	Lifetime short 1 nS
Ideal for amplifying short high power pulses	$\tau = 10 \text{ mS}$ Mean signal power governs amplifier performance	Will not amplify short high power pulses	$\tau = 1 \text{ nS}$
Poor switching properties (long response time)	$\tau = 10 \text{ mS}$	Good switching properties (fast switching)	$\tau = 1 \text{ nS}$
Technology for amplification in first and second windows emerging	New host glasses have to be developed for dopants	Amplification in 2nd telecommunications window easily achieved	SOA structure is based on mature laser technology

13.5 Gain Equalization

Optical amplifiers, like most electronic amplifiers, do not have a flat response curve over the entire range for which they have gain. The nonlinear response of EDFAAs has been theoretically modeled by Liu et al. [18]. Signal power, pump power, and frequency were all found to influence the linearity of the amplifier response. In the optical case, this variation in gain can be compensated by using optical techniques. The basic approach is to use gain flattening filters, fitted to the inverse of the gain spectral curve, to equalize the gain over the passband of the amplifier. The techniques that are more commonly used are summarized in Table 13.3.

When back reflecting filters are used, isolators may be required to protect light sources.

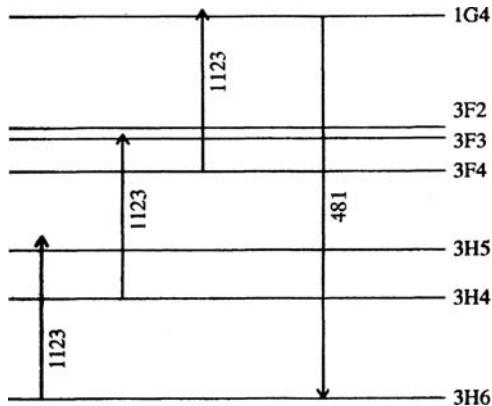
Table 13.3 Optical amplifier gain equalization methods

Thin film dielectric filters	(Hundreds of dielectric layers may be required. Light is selectively back reflected.)
Interferometer filters, such as Mach-Zehnder	(3 – 5 filters are required to cover the C-band for EDFA)
Long-period fiber gratings	(Grating density variations are formed in the core, with spacing $\Lambda = 200$ to 400 times wavelength. Selected light is forward coupled into cladding modes where it is attenuated.)
Chirped fiber Bragg gratings.	(Light is selectively back reflected.)

13.6 Fiber Lasers

In preceding sections of this chapter the characteristics of glass optical fiber amplifiers were discussed. In many cases the gain is sufficient to overcome losses so that lasing can be established at some wavelength. Fiber lasers have the advantage that they can be many meters long and losses are relatively low, allowing large output power to be obtained. For example, Paschotta et al. [19] have made a thulium-doped, blue-light laser that emits 230 mW at a wavelength of 481 nm. It is pumped by a 7-W diode laser at 807 nm that pumps a 1.6-W Nd: YAG at 1123 nm. There are three steps of upconversion leading ultimately to the emission of 481 nm photons, as shown in Fig. 13.9. The fiber was 2.2 m long. The output blue-light power increased linearly with the 1123 nm pump power up to an output of about 120 mW, then became slightly sublinear up to its maximum of 230 mW at a pump power of 1600 mW. In the linear range the slope efficiency was 18.5%. Some degradation of the output power was observed at the higher power levels, with the slope efficiency dropping to about 14% after about 20 min. However, operation at a 20 mW output level for an hour restored the slope efficiency to 18%. The authors suggested that modifying the fluoride glass composition to reduce color centers should enable more stable operation at high power levels.

Fig. 13.9 Energy levels of Tm³⁺ in fluoride glass [19]
© 1997 IEEE



Another advantage of fiber lasers is that they have an inherently narrow linewidth. This is true because the lasing transition are occurring between well defined energy levels rather than bands of states as in a semiconductor. This feature is demonstrated by a tunable, Tm³⁺-doped, fluoride fiber laser described by McAleavy et al. [20]. That laser emits at 2.3 μm wavelength with a linewidth of 207 MHz. The laser is 1.8 m long (including the external cavity) and has a threshold pump power of 31.4 mW, a slope efficiency of 19%, and a tuning range of 130 nm. Tuning is accomplished with a Littman/Metcalf external cavity composed of a diffraction grating and plane mirror. The narrow linewidth, coupled with tunability, make the laser suitable for detection of hydrocarbon gases, which have strong absorption lines in the spectral region spanned. For example, methane (CH₄) has a strong absorption centered at 2.3 μm.

Fiber lasers can also be used for generation of a train of high speed pulses by using the mechanism of mode-locked self oscillation. For example, Wu and Dutta [21] have made a harmonic mode-locked fiber laser based on a 23 m long erbium-doped fiber. A diagram of the laser is shown in Fig. 13.10. Both harmonic mode locking and rational harmonic mode locking were observed. For harmonic mode locking it is required that:

$$f_m = n f_c, \quad (13.12)$$

where f_m is the modulation frequency of the locking signal introduced by the intra-cavity modulator and f_c is the inverse of the cavity round-trip time (which is also the fundamental frequency of the fiber laser ring). Rational harmonic mode locking in an actively mode-locked laser occurs when the following condition is met:

$$f_m = (n + 1/p) f_c, \quad (13.13)$$

where n and P are both integers.

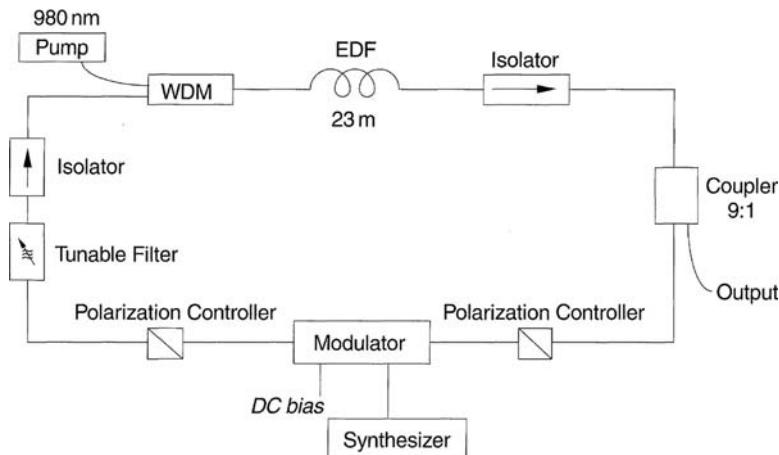


Fig. 13.10 Mode-locked fiber laser [21] © 2000 IEEE

In the laser of Fig. 13.10, the 22nd order of rational harmonic mode locking was observed when $f_m \cong 1$ GHz. An optical pulse train with a repetition frequency of 40 GHz was generated with a modulation frequency of $f_m = 10$ GHz. The locking signal was injected by a LiNbO₃ electrooptical modulator with a bandwidth of 11 GHz, and a wavelength-tunable optical filter with a 3-dB bandwidth of 1.0 nm. A synthesizer was used to generate the RF signal, which was amplified to 30 dBm to drive the modulator.

In Chapters 12 and 13 the basic characteristics of lasers and optical amplifiers have been discussed. In the next chapter we will return to the semiconductor laser to consider the very significant improvement in performance that results when a field-confining heterostructure is incorporated into the device.

Problems

- 13.1 (a) Name the three types of semiconductor optical amplifiers.
 (b) Explain how each of these differs from the others.
- 13.2 Which of the following wavelengths of light can be used to pump an erbium-doped fiber amplifier (EDFA)? 980 nm, 982 nm, 855 nm, 1350 nm, 1480 nm.
- 13.3 If an EDFA develops a stimulated emission power of 900 mW in response to an input signal power of 1 mW and the amplified spontaneous emission power is 50 mW, what is the optical gain of the amplifier in dB?
- 13.4 A fiber laser has a cavity length of 2 m and an effective index of refraction of 1.75. It emits at a vacuum wavelength of 2.4 μm. If it is to be harmonically modelocked to the third harmonic of its fundamental frequency, what must be the modulation frequency of the locking signal?

- 13.5 What are the ranges of wavelength over which each of the following types of optical amplifier can be used?
- semiconductor optical amplifier (SOA)
 - erbium doped fiber amplifier (EDFA)
 - Raman amplifier
 - Thulium doped fiber amplifier (TDFA)
- 13.6 A double-heterostructure laser diode is used as a Fabry-Perot-type optical amplifier. When biased at the operating point it has a gain coefficient of 50 cm^{-1} , a loss coefficient of 4 cm^{-1} and a confinement factor of 0.9. The length between Fabry-Perot endfaces is $350 \mu\text{m}$ and the reflectivities of the endfaces are both 33%.
- What is the overall gain of this optical amplifier?
 - If the reflectivities of both endfaces are made very small so that the device acts as a traveling wave amplifier, but the current through the device is kept to be the same as in (a), what is the overall gain?
- 13.7 If an EDFA operating at a wavelength of 1536 nm develops a stimulated emission power of 850 mW, how many photons/s are being generated?
- 13.8 What is the gain for a Fabry-Perot laser amplifier with both mirror reflectivities = 65% if the single-pass gain is 15?

References

- M. Faké, D.G. Parker: Optical Amplifiers, in *Photonic Devices and Systems*, (ed.) R.G. Hunsperger, Marcel Dekker, Ser. Opt. Eng., Vol. 45 (Marcel Dekker, New York, Basel 1994)
- C.K. Liu, F.S. Lal, J.J. Jou, M.C. Chang: Temperature and electromagnetic effects on erbium-doped fiber amplifier systems. *Opt. Eng.* **37**, 2095 (1998)
- M. Yamada, M. Shimizu, K. Yoshino, M. Horiguchi, M. Okuyasu: Temperature dependence of signal gain in erbium-doped fiber amplifiers *IEEE J. Quant. Electron.* **28**, 640 (1992)
- M. Islam, M. Nietubyc: Raman amplification opens the S-band window. WDM Solutions (A supplement to Laser Focus World) **37**, 53 (2001)
- S. Jiang, T. Luo, B.-C. Hwang, G. Nunzi-Conti, M. Myers, D. Rhonehouse, S. Honkanen, N. Peyghambarian: New Er³⁺-doped phosphate glass for ion-exchanged waveguide amplifiers. *Opt. Eng.* **37**, 3282 (1998)
- S. Honkanen, T. Ohtsuki, S. Jiang, S. Najafi, N. Peyghambarian: High Er concentration phosphate glasses for planar waveguide amplifiers. *Proc. SPIE* **2996**, 32 (1997)
- T. Kitagawa, K. Hattori, K. Shuto, M. Yasu, M. Kobayashi, M. Horoguchi: Amplification in erbium-doped silica-based planar light-wave circuits. *Electron. Lett.* **28**, 1818 (1992)
- R.N. Ghosh, J. Shmulovich, C.F. Kane, M.R.X. Barros, G. Nykolak, A.J. Bruce, P.C. Becher: 8 mW threshold Er³⁺-doped planar waveguide amplifier. *IEEE Photon. Technol. Lett.* **8**, 518 (1996)
- S. Jiang, J.D. Myers, D. Rhonehouse, M. Myers, R. Belford, S. Hamlin: Laser and thermal performance of a new erbium doped phosphate laser glass. *Proc. SPIE* **2138**, 166 (1994)
- S.I. Najafi: Overview of Nd- and Er-doped glass integrated optics amplifiers and lasers. *Proc. SPIE* **2996**, 54 (1997)

11. V.P. Gapontsev, S.M. Matitsin, A.A. Isineer. V.B. Kravchenko: Erbium glass lasers and their applications. *Opt. Laser Technol.* **14**, 189 (1982)
12. J. Buns, R. Plastow: *IEEE J. QE-21*, 614 (1985)
13. S. Kobayashi, T. Kimura: *IEEE Spectrum* **21**, 26–33 (May 1984)
14. Y.C. Zhang, Z.X. Qin, S.L. Wu, L.J. Wu, L.J. Wang, D.E. Lee: *Fiber Integrated Opt.* **8**, 99 (1989)
15. T. Brenner, R. Dall, A. Holtmann, P. Besse, N. Melchior: *IEEE 5th Int'l Conf. on InP and Related Materials*, Paris (1993) Digest p. 88
16. J.E. Johnson, L.J.-P. Ketselsen, J.A. Grenko, S.K. Sputz, J. Vandenberg, M.W. Focht, D.V. Stampone, L.J. Petricolas, L.E. Smith, K.G. Glogovsky, G.J. Przybylek, S.N.G. Chu, J.L. Lentz, N.N. Tzafaras, L.C. Luther, T.L. Pernell, E.S. Walters, D.M. Romero, J.M. Freund, C.L. Reynolds, L.A. Gruezke, R. People, M.A. Alam: Monolithically integrated semiconductor optical amplifier and electroabsorption modulator with dual-waveguide spot-size converter input. *IEEE J. Select. Topics Quan. Electron.* **6**, 19 (2000)
17. H. Hatakeyama, T. Tamanuki, K. Moriea, T. Sasaki, M. Yamaguchi: Uniform and high-performance eight-channel bent waveguide SOA array for hybrid PICs. *IEEE Phot. Tech. Lett.* **13**, 418 (2001)
18. C.K. Liu, F.S. Lai, J.J. Jou: Analysis of nonlinear response in erbium doped fiber amplifiers. *Opt. Eng.* **39**, 418 (2001)
19. R. Paschotta, N. Moore, W.A. Clarkson, A.C. Tropper, D.C. Hanna, G. Maze: 250 mW of blue light from a thulium-doped upconversion fiber laser *J. Select. Topics Quant. Electron.* **3**, 1100 (1997)
20. F.J. McAlevey, J. O'Gorman, J.F. Donegan, B.D. MacGraith, J. Hegarty, G. Maze: Narrow linewidth, tunable Tm^{3+} -doped fluoride fiber laser for optical-based hydro-carbon gas sensing. *IEEE J. Select. topics Quant. Electron.* **3**, 1103 (1997)
21. C. Wu, N.K. Dutta: High-repetition-rate optical pulse generation using a rational harmonic mode-locked fiber laser. *IEEE J. Quant. Electron.* **36**, 145 (2000)

Supplementary Reading on Optical Amplifiers

- M.J. Connelly: *Semiconductor Optical Amplifiers* (Springer, New York, 2002)
- N.K. Dutta, Q. Wang: *Semiconductor Optical Amplifiers* (World Scientific Publishing Company, Hackensack, NJ, 2006)
- M.J.F. Digonnet: *Rare-Earth-Doped Fiber Lasers and Amplifiers, Second Edition, Revised and Expanded* (Marcel Dekker, New York, 2001)
- A. Bjarklev *Optical Fiber Amplifiers: Design and System Applications* (Artech House, Norwood, MA, 1993)

Chapter 14

Heterostructure, Confined-Field Lasers

In Chapter 12, it was demonstrated that confining the optical field to the region of the laser in which the inverted population exists results in a substantial reduction of threshold current density and a corresponding increase in efficiency. As early as 1963, it was proposed that heterojunctions could be used to produce a waveguiding structure with the desired property of optical confinement [1, 2]. At about the same time, others proposed using a heterojunction laser structure not for optical field confinement, but to produce higher carrier injection efficiency at the p-n junction, and to confine the carriers to the junction region [3, 4]. Actually, all three of these mechanisms are present in a heterostructure laser, and their combined effects result in a device that is vastly superior to the basic p-n homojunction laser.

Because of the technological difficulties involved in the growth of multilayer heterostructures, several years passed before operational heterostructure lasers were fabricated in 1969 [5–7]. These first devices were all fabricated in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ because the close lattice match between GaAs and AlAs, which resulted in minimal strain at the interfaces between the heterostructure layers, made that material preferable to others, such as GaAsP. Even in these early heterostructure lasers, threshold current densities were on the order of 10^3 A/cm^2 rather than 10^4 to 10^5 A/cm^2 , as in a comparable homojunction laser.

Hayashi et al. [5], and Kressel and Nelson [6] fabricated single-heterojunction (SH) lasers in 1969, while Alferov et al. [7] fabricated more effective double-heterojunction (DH) lasers. Since 1969, numerous improvements have been made in the basic heterostructure laser, resulting in devices with threshold current densities on the order of 10^2 A/cm^2 . The heterojunction laser has become the standard light source in OIC's and in fiber-optic communications systems. For their work in developing this device, Zhores Alferov and Herbert Kroemer were jointly awarded the Nobel Prize in Physics for 2000, along with Jack Kilby for his part in the invention of the integrated circuit.

In this chapter, both the basic and the advanced heterojunction laser structures are discussed. The relationships between geometric and materials properties, and the performance characteristics of the laser are described. In Section 14.5 of this chapter, we also consider the important question of laser diode reliability. The DH injection laser, while very efficient, is one of the most highly-stressed semiconductor

devices in terms of intense optical and electrical fields. Hence a great deal of effort has been expended to produce devices with satisfactory lifetime and limited degradation of performance.

14.1 Basic Heterojunction Laser Structures

14.1.1 Single Heterojunction (SH) Lasers

The simplest heterojunction laser to fabricate is the SH structure shown in Fig. 14.1 [5, 6]. In the fabrication of this device, the anomalously fast diffusion of Zn in GaAs [8, 9] is utilized to form a diffused p-n junction lying 1–2 μm below the $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ -GaAs heterojunction. If the n- and p-type dopant concentrations are approximately equal on both sides of the p-n junction, the injection current will consist mostly of electrons injected into the p-type layer, because the effective mass of an electron is about 7 times less than that of a hole in $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ [10]. Thus, the inverted-population region, or active layer, in this type of SH laser is in the p-type GaAs, as shown in Fig. 14.1. This SH laser can be fabricated by using the method of liquid epitaxial growth that is described in Chapter 4, except that a relatively high growth temperature of 900°C–1000°C is used in order to promote the required Zn diffusion into the substrate.

The thickness of the active layer can be selected by controlling the time and temperature of epitaxial growth (and hence Zn diffusion) to produce a thickness of about 1–5 μm . However, since the diffusion length of injected electrons is only about 1 μm , increasing the thickness of the p-GaAs layer beyond that value results in decreased efficiency and higher threshold current density, because the inverted population region is still limited to a thickness of about 1 μm by electron recombination [11]. Thus, although the optical mode spreads over the entire p-GaAs layer, it can be pumped by stimulated emission over only a 1 μm -thick layer closest to the p-n junction, and reduced efficiency results. In some case, it might be desirable to increase the p-GaAs layer thickness to greater than 1 μm , even at the expense of increased threshold current density, because reduced diffraction of the optical beam

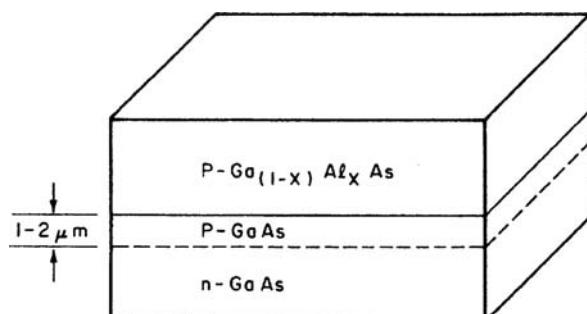


Fig. 14.1 Single heterostructure laser diode

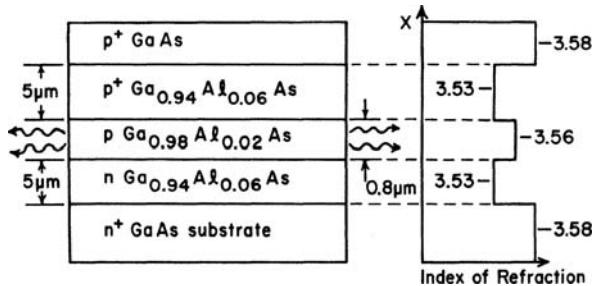
in the thicker layer results in a smaller divergence angle of emitted light in the plane perpendicular to the p-n junction.

In the SH laser, optical confinement occurs only on one side of the light-emitting junction, at the interface between the p-GaAs and p-Ga_(1-x)Al_xAs layers. Although there is a waveguiding effect in the depletion layer of the p-n junction itself because of carrier concentration reduction [12], that phenomenon is generally negligible compared to the substantial confinement that occurs because of the change in index of refraction at the heterojunction. Thus the SH laser structure is only partially effective in producing the desired optical confinement. As a result, SH lasers exhibit higher threshold current densities than do comparable DH lasers. In fact, SH lasers must be operated on a pulsed basis, rather than cw, at room temperature. In many applications, pulsed operation is not detrimental to overall system performance, and may even be beneficial, to improve signal to noise ratio in signal processing. Hence SH lasers have been widely used in the past as sources in a variety of opto-electronic applications. However, when a cw laser diode source operating at room temperature is required, a double heterostructure laser must be used. In recent years the large demand for double heterostructure lasers has resulted in large-volume production and lowered cost, so SH lasers are rarely used.

14.1.2 Double Heterostructure (DH) Lasers

The physical structure of a typical DH laser in GaAlAs is shown in Fig. 14.2, along with a diagram of the index of refraction profile in the direction normal to the p-n junction plane. The basic GaAlAs, three-layer waveguide structure is usually grown on a heavily-doped n⁺-type substrate, and is capped with a heavily-doped layer of p⁺-GaAs, to facilitate the formation of electrical contacts. The active region is established on the p-type side of the p-n junction, as explained previously in Section 14.1.1. Very often, the active layer will contain a certain concentration of Al in order to shift the optical emission to shorter wavelengths. This topic is discussed in more detail in Section 14.3. The thickness of the active layer should be less than 1 μm in order to assure that an inverted population exists throughout the entire layer rather than being limited by the diffusion length for injected electrons. In fact, active layer thickness is often reduced to 0.2–0.3 μm in order to produce greater population inversion and laser photon density. Typical electrical dopant concentrations are $N_A \cong 2 \times 10^{19} \text{ cm}^{-3}$ for the p⁺-layer, $N_A = 1 \times 10^{16} \text{ cm}^{-3}$ for the active layer, $N_D = 1 \times 10^{17} \text{ cm}^{-3}$ in the n-GaAlAs, and $N_D = 2 \times 10^{18} \text{ cm}^{-3}$ in the substrate. The choice of these concentrations is governed by the desire to reduce series resistance in the bulk material, while limiting free-carrier absorption in the light-emitting region. The multilayer Ga_(1-x)Al_xAs structure shown in Fig. 14.2 is often grown by the slidebar method of liquid epitaxial growth that is described in Chapter 4. For a detailed description of the application of this method to DH laser fabrication see Casey and Panish [13]. The newer methods of MOCVD or MBE growth (also described in Chap. 4) offer some advantages for laser diode fabrication. MOCVD permits the processing of relatively large wafers, while quantum well

Fig.14.2 Double heterostructure laser diode



lasers can be made by MBE or MOCVD. Therefore these methods are now replacing slidebar LPE in many production facilities. For a review of the development of heterojunction lasers and the fabrication methods used to produce them see Kroemer [14], and Iga and Kinoshita [15].

The double heterojunction laser structure provides confinement on both sides of the active region. Because of this confinement of both lasing mode photons and injected carriers to the inverted population region, in which gain due to stimulated emission is possible, the DH laser is highly efficient, and requires minimal threshold current, as compared to other semiconductor lasers.

14.2 Performance Characteristics of the Heterojunction Laser

The superior performance of heterojunction lasers results from the combined effects of optical field confinement and more efficient carrier injection and recombination. In this section these basic features of the phenomena are described, and the differences between the performance of DH, SH and homojunction lasers are explained.

14.2.1 Optical Field Confinement

In Chapter 12, it was shown that the optical field confinement, as characterized by the fraction of lasing mode photons that remain in the inverted population region d/D , strongly affects both threshold current density J_{th} and efficiency η . The quantities J_{th} and η can be calculated using (12.10) and (12.17), but one must know the correct value of d/D to use in any given case.

The procedure for determining d/D from first principles is to solve the wave equation, subject to the boundary conditions appropriate for the particular waveguide structure, thereby determining the quantitative expression for the mode shape. The ratio d/D is then found by integrating the expression for photon density over the thickness of the inverted population layer and dividing that quantity by the total number of photons in the mode, which is obtained by integrating the photon density expression over the extent of the entire mode. The solution of the wave equation to find the mode shape in a three-layer, symmetric waveguide, as described in

Chapter 3, is a lengthy problem, involving extensive computer calculations. However, in many cases, sufficient accuracy can be obtained by using an approximate, but more concise, set of relations developed by McWhorter [16].

By solving Maxwell's equations, *McWhorter* has obtained relations that give the transverse mode shape in a semiconductor laser for an active region of thickness d and index n_a , bounded by confining regions with indices n_b and n_c , as shown in Fig. 14.3, such that n_a is greater than n_b and n_c . It is assumed that the light confining layers are sufficiently thick so that the tails of the optical mode do not penetrate through to the p⁺ and n⁺ contact layers. The relative spatial energy density Φ (photon density) in the active region (a) is given by

$$\Phi = A \cos k_a x + B \sin k_a x \quad \text{for} \quad -d/2 \leq x \leq d/2 , \quad (14.1)$$

where A and B are constants.

In the confining layers (b) and (c) Φ is given by

$$\Phi = e^{-k_b x} \quad \text{for} \quad x > d/2 , \quad (14.2)$$

and

$$\Phi = e^{+k_c x} \quad \text{for} \quad x < -d/2 . \quad (14.3)$$

The extinction coefficients k_b and k_c are determined from

$$k_b = \left(\frac{k_0^2 d}{2} \right) \left(\frac{n_a^2 - \bar{n}^2}{n^2} \right) + \frac{1}{2d} \left(\frac{n_c^2 - n_b^2}{n_a^2 - \bar{n}^2} \right) , \quad (14.4)$$

and

$$k_c = \left(\frac{k_0^2 d}{2} \right) \left(\frac{n_a^2 - \bar{n}^2}{n^2} \right) + \frac{1}{2d} \left(\frac{n_b^2 - n_c^2}{n_a^2 - \bar{n}^2} \right) , \quad (14.5)$$

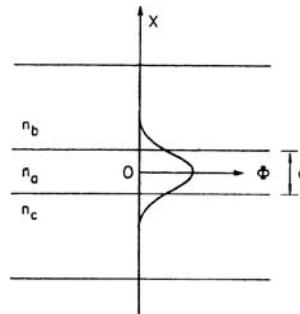


Fig. 14.3 Model for field confinement in a double hetero-structure laser

where n is the index of refraction of pure GaAs at the lasing wavelength, and where

$$\overline{n^2} = \frac{n_b^2 + n_c^2}{2}. \quad (14.6)$$

and

$$k_0 = 2\pi/\lambda_0. \quad (14.7)$$

Once k_b and k_c have been calculated, the constants k_a , A and B can be determined from (14.1), (14.2) and (14.3) by matching boundary conditions for continuity at the interfaces in the usual way. Then the confinement factor d/D can be determined from

$$\frac{d}{D} = \frac{\int_{-d/2}^{d/2} \Phi dx}{\int_{-\infty}^{\infty} \Phi dx}, \quad (14.8)$$

where Φ is taken from (14.1), (14.2) and (14.3).

The calculation of d/D from McWhorter's relations is still a lengthy procedure, even though it is more brief than the direct solution of Maxwell's wave equation. Fortunately, another alternative exists. Casey and Panish [17] have developed an approximate expression for the confinement factor d/D that holds for the special case of a symmetric, three-layer, $\text{Ga}_{(1-a)}\text{Al}_a\text{As}-\text{GaAs}-\text{Ga}_{(1-a)}\text{Al}_a\text{As}$ waveguide with small thickness d . They have shown that, for cases in which

$$d \lesssim \frac{0.07\lambda_0}{a^{1/2}}, \quad (14.9)$$

where a is the atomic fraction of Al in the confining layers, the confinement factor is given by

$$\frac{d}{D} \cong \frac{\int_0^{d/2} E_0^2 \exp(-2\gamma x) dx}{\int_0^{\infty} E_0^2 \exp(-2\gamma x) dx}, \quad (14.10)$$

where E_0 in the peak field amplitude and γ is given by

$$\gamma \cong \frac{(n_a^2 - n_c^2)k_0^2 d}{2} \quad (14.11)$$

where n_a and n_c are the indices of refraction in the active and confining layers, respectively. The confinement factor for a $\text{Ga}_{(1-a)}\text{Al}_a\text{As}-\text{GaAs}$ three-layer symmetric waveguide is plotted in Fig.14.4 for various typical values of d and a . These curves are based on the assumption that the laser is operating in the fundamental (TE_0) mode, which is generally the case in a well-designed DH laser. It can be seen from the data of Fig.14.4 that nearly 100% confinement can be obtained for active

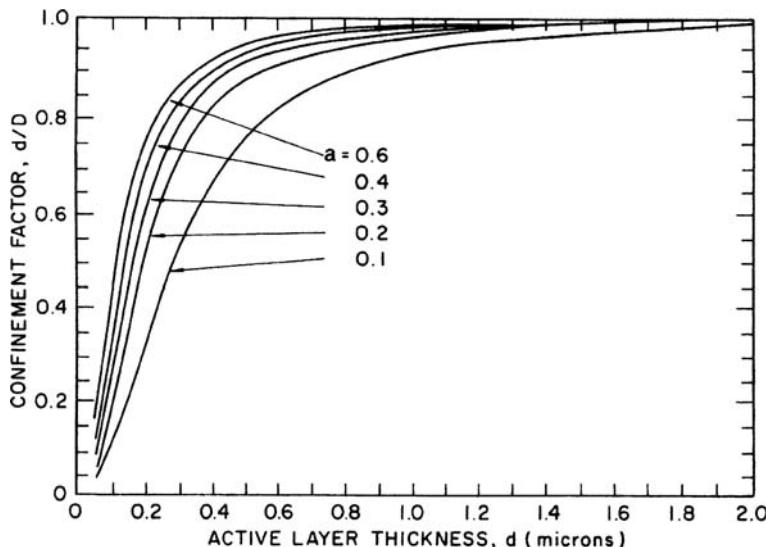


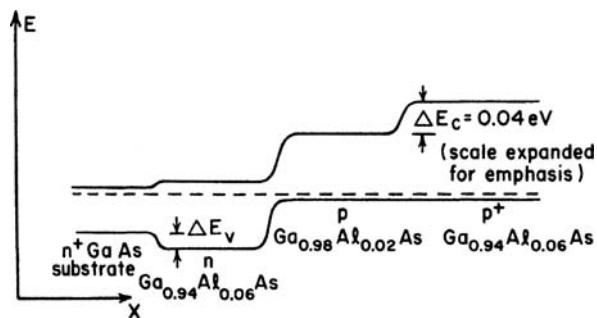
Fig. 14.4 Confinement factor for the fundamental mode in a symmetric, three-layer $\text{Ga}_{(1-a)}\text{Al}_a\text{As}$ - GaAs waveguide [17]

region thickness as small as $0.4 \mu\text{m}$, without exceeding an atomic fraction of Al equal to 0.6. This is significant because AlAs is hygroscopic, and larger Al concentrations in the $\text{Ga}_{(1-a)}\text{Al}_a\text{As}$ are therefore unadvisable. Finally, it should be noted that the confinement factor d/D can also be calculated by using commercially available optical-waveguide-mode modeling software that was mentioned in Chapter 3.

14.2.2 Carrier Confinement

As mentioned previously, some of the researchers who first proposed the heterojunction laser [3, 4] did so not because of optical field confinement, but rather because the heterojunction offers improved carrier injection efficiency and carrier confinement to the active region. The energy band diagram of a $n^+ - n - p - p^+$ GaAlAs double heterojunction laser is shown in Fig. 14.5. Because the bandgap is larger in the regions with larger Al concentration, discontinuities result in the conduction band at the $p - p^+$ junction (ΔE_c) and in the valence band at the $n - p$ and $n^+ - n$ junctions (ΔE_v). The energy bands are shown in Fig. 14.5 for the case of no applied external bias voltage, and the magnitude of the conduction and valence band discontinuities has been drawn larger than to scale in order to emphasize their presence for pedagogic purposes. When a forward bias voltage is applied to the DH laser, electrons are injected from the n -region into the p -region, forming the desired recombination current. The conduction band discontinuity at the $p - p^+$ interface ΔE_c provides a barrier to the injected electrons, tending to confine them to the p -region, and to enhance the

Fig. 14.5 Energy band diagram of a n^+ -n-p-p $^+$ double heterojunction laser diode. Zero bias state is assumed, i.e. no externally applied voltage is present



probability of their recombining with holes by the stimulated emission process. The valence band discontinuity at the n-p junction ΔE_v augments the already existing built-in potential barrier, to further prevent hole injection into the n region, thereby improving injection efficiency. Thus the double heterostructure tends to confine both majority and injected minority carriers to the active p-layer. Since the photons of the optical mode are also confined to the active layer by the heterojunctions as described in Section 14.2.1, the DH laser provides optimum conditions for establishing both the largest possible inverted population and the greatest photon density in the active layer. It will be recalled from Chapter 11 that these are the two primary requirements for stimulated emission. Thus the DH laser can be expected to significantly outperform the homojunction laser, and indeed it does.

14.2.3 Comparison of Laser Emission Characteristics

Threshold current density for a typical homojunction laser is on the order of 10^4 A/cm 2 , and differential quantum efficiency (above threshold) is about 10%. Even though the SH laser provides photon and carrier confinement at only one interface of the active layer, it is effective enough to reduce the threshold current density to about 5×10^3 A/cm 2 , and to increase differential quantum efficiency to about 40%. It must be remembered, of course, that J_{th} and η_D are functions of temperature, active layer thickness, and dopant concentrations in any given case; so the numbers quoted here are typical values provided for comparison purposes only. Because J_{th} is still relatively large in SH lasers and η_q is only moderate, they cannot be operated on a cw basis. They must be pulsed, usually with 100 ns length pulses at a repetition rate of 100–1000 Hz, to allow time for the junction region to cool between pulses. Nevertheless, peak pulse output power from 10–30 W is possible.

The DH laser, which provides confinement at both interfaces of the active region, has a threshold current density typically from 100 to 400 A/cm 2 , and differential quantum efficiency as high as 91% [18]. The DH laser can be operated cw, with output power as high as several Watts [19]. By constructing bars of semiconductor diode lasers in line Schulz and Poprawe [20] have achieved an output as high

as 267 W/bar. One disadvantage of the DH laser is that the angular divergence of the beam in the plane perpendicular to the junction can be 20–40°, rather than the 15–20° commonly exhibited by SH lasers. This divergence, which results from diffraction in the thin active layer, can be reduced by using special large optical cavity (LOC) DH structures, which provide separate confinement of photons and carriers. For broad area homojunction, SH and DH lasers, such as have been described so far, there is no lateral confinement in the plane of the junction, light is free to spread over the full 20–80 μm width of the chip. As a result, beam divergence in the junction plane is only about 10° in all cases. Stripe geometry lasers, which are described in Section 14.4.1, also confine the light in the lateral direction. This confinement results in a further reduction in J_{th} , but it also causes an increase in beam divergence in the junction plane. Surface-emitting lasers, to be described in Section 14.6, have a large emitting area, which results in relatively small beam divergence.

14.3 Control of Emitted Wavelength

One of the more important characteristics of semiconductor lasers is the wavelength at which the peak emission of light occurs. It is often desirable to shift the peak wavelength somewhat from that which is characteristic of the substrate semiconductor material in order to take advantage of waveguide transparency in a given range.

14.3.1 $Ga_{(1-x)}Al_xAs$ Lasers for Fiber-Optic Applications

The peak emission wavelength of a homojunction laser formed in GaAs is about 9200 Å, corresponding to radiative recombination of electrons in donor states 0.005 eV below the conduction band edge and holes in acceptor states 0.030 eV above the valence bandedge, (The bandgap of pure GaAs at room temperature is 1.38 eV [21]) While GaAs lasers can be used as light sources with glass optical fiber waveguides, their performance is less than optimum because the fibers can have significant absorption loss at 9200 Å. The spectral attenuation curve for a moderately pure borosilicate glass fiber, typical of those available commercially, is shown in Fig.14.6. The absorption peak at about 9400 Å is caused by the presence of OH ions in the glass. In order to avoid this absorption, $Ga_{(1-x)}Al_xAs$ heterojunction lasers can be used, with sufficient Al added to the active region to shift the emitted wavelength to 8500 Å. The emission wavelength of $Ga_{(1-x)}Al_xAs$ can be shifted in the range from 9200 Å to about 7000 Å by the addition of Al, as shown in Fig.14.7. Shorter wavelengths are unattainable because the band gap of the $Ga_{(1-x)}Al_xAs$ becomes indirect for Al concentration greater than about 35% and as a result internal quantum efficiency is greatly reduced. Because of their compatibility with readily available and inexpensive optical fibers, $Ga_{(1-x)}Al_xAs$ heterostructure lasers, emitting at 8500 Å are widely used today in local area networks and other short-distance applications.

Fig. 14.6 Typical spectral attenuation curve for commercially available multimode glass fiber

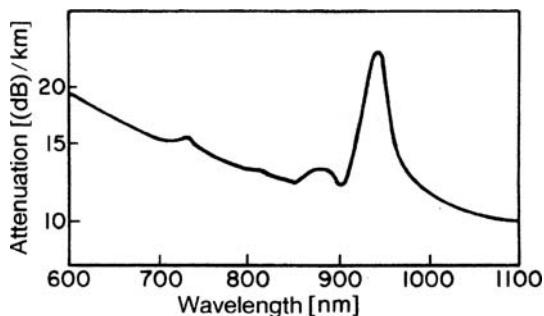
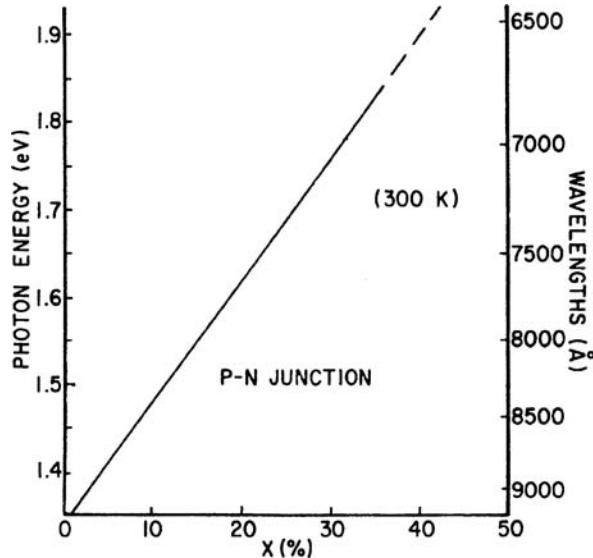


Fig. 14.7 Peak emission wavelength of a $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ laser as a function of Al concentration in the active region. X is the atomic fraction of Al



For long-distance telecommunications the GaAlAs laser has been replaced by GaInAsP laser diodes operating at either 1.3 or 1.55 μm . During the 1980s, the equality of optical fibers continuously improved, due largely to better purification. The OH ions can be removed, along with other contaminants, to produce optical fibers with minimal attenuation, approaching the limit imposed by Rayleigh scattering. The spectral attenuation curve for one such fiber is shown in Fig. 14.8. The minima in attenuation that occur at 1.3 μm and 1.55 μm are particularly important because material dispersion for a silica-rich core is theoretically equal to zero at a wavelength of 1.27 μm [22] and can be shifted to yield zero dispersion at 1.55 μm . Thus, operation at those wavelengths yields not only minimum attenuation, but also minimum dispersion. Of course, in order to obtain minimum dispersion, either a single-mode or a graded index fiber must be used so that modal

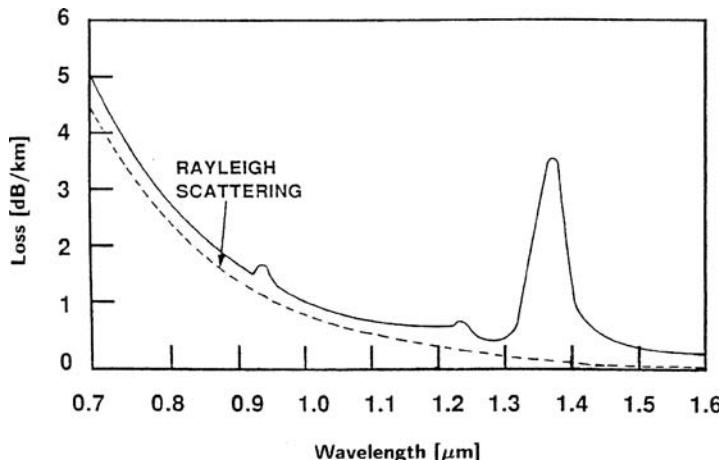


Fig. 14.8 Attenuation in a highly-purified step index single-mode fiber consisting of a borosilicate cladding surrounding a phosphosilicate core

dispersion is avoided. The details of optical fiber design are beyond the scope of this book, but are described elsewhere [23].

14.3.2 Lasers Made of Quaternary Materials

As described in the previous section, there is a distinct advantage to using a wavelength of about $1.3 \mu\text{m}$ or $1.55 \mu\text{m}$ in a fiber-optic system. However, $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ lasers cannot emit light at these wavelengths. Because of this limitation, GaInAsP has become the preferred material for fabricating lasers for use in long-distance telecommunications and data transmission. As shown in Fig. 4.12, the bandgap of GaInAsP can be varied from 1.4 eV to 0.8 eV , corresponding to emission wavelengths of $0.886 \mu\text{m}$ to $1.55 \mu\text{m}$. Lattice mismatch at the interfaces between the various layers of the heterostructure, which would result in defect centers that greatly increase absorption and non-radiative recombination, can be avoided by proper choice of the concentrations of the constituent elements. GaInAsP has been demonstrated to be an effective material for producing lasers operating at both $1.3 \mu\text{m}$ and $1.55 \mu\text{m}$ wavelengths [24].

14.3.3 Long-Wavelength Lasers

Semiconductor lasers operating in the $2\text{--}5 \mu\text{m}$ wavelength range are of considerable interest for future optical communications systems employing low-loss fluoride glass fibers, and for chemical gas analysis and atmospheric pollution monitoring. Double-heterostructure (DH) lasers in various III-V materials have been used to achieve emission in this wavelength range. Mani et al. [25] have reported DH

lasers prepared by LPE in the InAsSbP/InAsbP system, on InAs substrates. These lasers emit at $3.2\text{ }\mu\text{m}$, with a pulsed threshold current density of 4.5 KA/cm^2 at 78 K . Martinelli [26] has produced DH lasers emitting at $2.55\text{ }\mu\text{m}$ in InGaAsP/InP. Threshold currents of 650 A/cm^2 have been achieved at 80 K ; and output powers of several milliwatts were typical.

Rare-earth-doped II-VI materials have also been utilized to produce long-wavelength emitting lasers. For example, Ebe et al. [27] have obtained emission at wavelengths between 4.0 and $5.5\text{ }\mu\text{m}$ at 77 K in PbEuTe DH lasers. Threshold current was approximately 0.5 kA/cm^2 . These devices could also be operated at over 200 K ambient temperature with thermoelectric cooling. GaInAsSb and PbSe laser diodes have been reported by Sorokina and Vodopyanov [28].

14.4 Advanced Heterojunction Laser Structures

14.4.1 Stripe Geometry Lasers

In the discussions of semiconductor lasers thus far, it has been tacitly assumed that the width of the laser was much larger than the thickness of the active layer, so that optical confinement occurred in only one of the transverse directions, i.e. perpendicular to the junction plane. If the width of the laser is limited in some fashion to produce a narrow stripe geometry of typically $5\text{--}25\text{ }\mu\text{m}$ width, the device will exhibit altered performance characteristics. The most important of these is that the threshold current will be reduced because of the smaller cross sectional area available for current flow. If the width of the stripe is about $10\text{ }\mu\text{m}$ (or less), the lateral confinement will also result in fundamental TE_{00} mode operation, as discussed in more detail in Section 14.4.2. Passivation of the longitudinal junction edges by removing them from the side surface of the laser chip reduces leakage current and lengthens the useful life of the laser by reducing degradation.

The lateral edges of the laser may be defined by masked etching of an oxide layer [29], as shown in Fig. 14.9a, or by masked proton bombardment [30], to produce high resistivity semi-insulating regions on either side of the stripe, as shown in Fig. 14.9b. A planar stripe laser can also be formed by masked diffusion [31], as shown in Fig. 14.9c, or by masked etching and epitaxial growth, as shown in Fig. 14.9d [32]. The lasers of Fig. 14.9a,b and c are described as being “gain-guided” lasers, because of the change in effective index in the lateral direction is produced by confining the current flow, and hence gain, to a central region. The laser of Fig. 14.9d is called an “index-guided” laser because the index is directly varied in the lateral direction by changing the chemical composition of the material.

14.4.2 Single-Mode Lasers

Most wide-cavity semiconductor lasers oscillate with a multiplicity of transverse and longitudinal modes present, as described in Chapter 12. In many applications

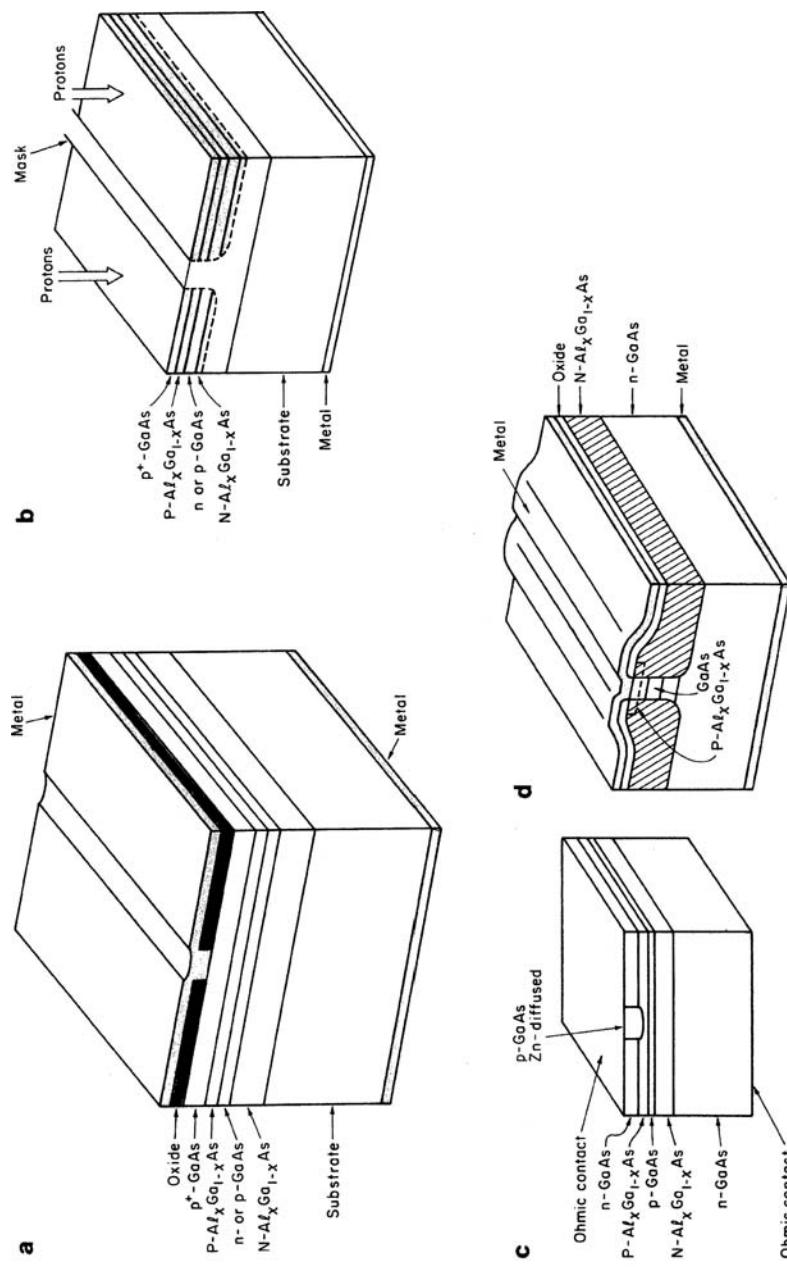


Fig. 14.9 a–d Stripe Geometry lasers. a Oxide insulating layer [29]; b proton bombarded insulating layer [30]; c planar diffused structure [31]; d epitaxially grown buried heterostructure [32]

this is not objectionable. However, when a high degree of phase coherence, or minimal dispersion, is required, as in the case of long-distance optical fiber communications, single-mode lasers become a necessity. Fortunately, the stripe-geometry gain-guided laser structures described in Section 14.4.1 are generally effective in establishing not only a single transverse mode, but a single longitudinal mode as well, as long as the stripe width is less than about twice the carrier diffusion length (approximately 3 μm in GaAlAs) [33]. Lasers with broader stripes, up to 29 μm in width, will also oscillate in a single mode. However mode instability is often observed, accompanied by *kinks* in the output curve of optical power versus current density, as shown in Fig. 14.10. Such kinks are undesirable in themselves, aside from any consideration of mode instability, because they prevent linear modulation of the laser light output. Mode instability and current kinks are thought to arise in relatively wide stripe-geometry lasers because of localized reductions the carrier density profile (*hole burning*) caused by carrier consumption at the peak of the optical mode profile [34, 35]. If the width of the stripe is made less than about two minority carrier diffusion lengths, carriers can diffuse from the edges of the stripe, where optical intensity is low, to replenish those consumed in the stimulated emission process. Hence, mode instability and current kinks are not observed in such lasers [33].

Stable single-mode lasers can be produced by gain guiding as described above or by fabricating a two dimensional structural waveguide, relying on index of refraction differences confine the optical mode. For example, in the stripe geometry lasers shown in Fig. 14.9a,b,c, optical confinement in the transverse direction is produced by the variation of the imaginary part of the dielectric constant, which is related to the index of refraction [35]. Hence, hole burning, which causes localized gain saturation, strongly affects optical confinement. In lasers that possess a two-dimensional structural waveguide, such as the buried heterostructure laser of Fig. 14.9d, gain saturation does not affect optical confinement. Hence stable single-mode oscillation can be obtained for stripe widths in excess of twice the carrier diffusion length

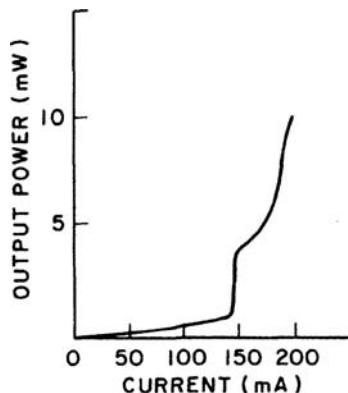


Fig. 14.10 Light output-current characteristic of a stripe-geometry laser with mode instability

[37, 38]. In addition to the buried heterostructure, other two-dimensional waveguide structures can be used to produce single-mode lasers. For a thorough discussion of the various techniques of mode control in semiconductor lasers see Kaminow and Tucker [39]. A survey of various specialized heterostructures and their operating characteristics has been published by Baets [40].

14.4.3 Integrated Laser Structures

The laser diode structures that have been described in preceding sections are generally suitable for monolithic integration with optical waveguides and other elements of an optical integrated circuit with regard to their dimensions and material composition. However, four significant problems must be dealt with in the design of any specific monolithically integrated laser/waveguide structure. There must be efficient coupling of the light from the laser to the guide, and there must be some means of providing the optical feedback required by the laser. Also, the absorption loss of the waveguide must be low at the emitted wavelength. Finally, some provision must be made for the electrical contacts. One structure that provides a solution to these problems, demonstrated by Hurwitz, is shown in Fig. 14.11. The active region of the laser is doped with Si acceptor atoms, so as to shift the emission wavelength of the GaAs to a $1 \mu\text{m}$ wavelength that is not strongly absorbed ($\alpha < 1 \text{ cm}^{-1}$) in the n-GaAs waveguide. The rectangular laser mesa, $300 \mu\text{m}$ long and $45\text{--}90 \mu\text{m}$ wide, is produced by masked chemical etching down through the epitaxially grown GaAlAs layers to the substrate. Orientation of the rectangular mask along the $<110>$ cleavage planes that are perpendicular to the $<100>$ wafer surface results in reasonably parallel end faces to form the mirrors of the Fabry-Perot cavity. The layers of SiO_2 coating the end faces are necessary because the index difference at the laser-waveguide interface would otherwise be insufficient to provide adequate reflection. After SiO_2 deposition, the GaAs waveguide was grown by vapor phase epitaxy. Electrical contacts were made to the p⁺ top surface and the n⁺ substrate. Room temperature threshold current densities as low as $7.5 \times 10^3 \text{ A/cm}^2$ were observed for these lasers.

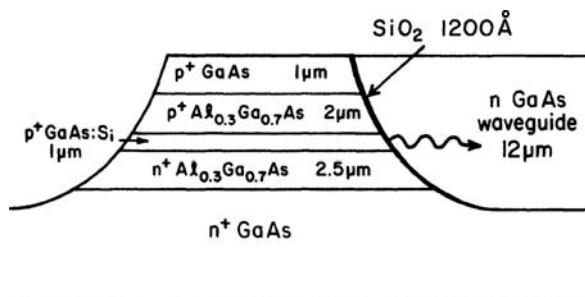
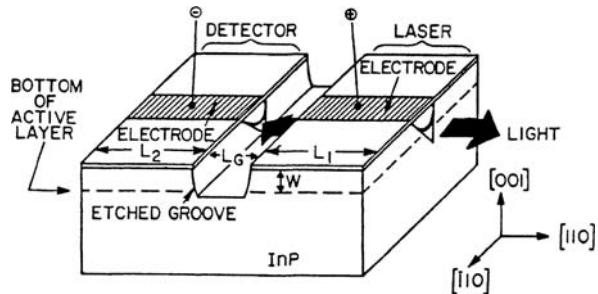


Fig. 14.11 A monolithic laser/waveguide structure in GaAlAs [41]

Fig. 14.12 A monolithically integrated GaInAsP laser and monitor photodiode [42]



Masked chemical etching of a laser facet has also been used by Koszi et al. [42] to produce a GaInAsP laser automatically integrated with a monitor photodetector as shown in Fig. 14.12. The laser in that case was a buried-heterostructure, stripe-geometry device emitting at a wavelength of $1.3\text{ }\mu\text{m}$. The isolation groove between the laser and photodiode was produced by masking with SiO_2 , configured by standard photolithography, and etching with a methyl alcohol and bromine solution. The threshold current and external quantum efficiency of the laser were measured to be in the range $I_{\text{th}} = 25 - 35\text{ mA}$ and $\eta_q = 0.06 - 0.09\text{ mW/mA}$. Comparable values for similar lasers with cleaved facets were $I_{\text{th}} = 20\text{ mA}$ and $\eta_q = 0.19\text{ mW/mA}$. The decreased performance of the lasers with etched facets was attributed to a slight tilt in the laser facet. Nevertheless, the I_{th} and η_q of the etched-facet lasers were reasonable for use in optical integrated circuits. A photocurrent of 15 nA per mW of laser optical output power was obtained from the photodiode, and the response was linear. Since light emission was into the air absorption loss was not a problem.

A technique for producing monolithically integrated laser diodes with cleaved facets has been developed by Antreasyn et al. [43, 44]. This method, called “stop cleaving”, is illustrated in Fig. 14.13, where the thick dark line marks the location of the laser cavity. Holes etched into the substrate are employed to stop the propagation along a cleavage plane, producing cleaved facets which are of comparable quality to conventionally cleaved laser facets. To produce stop-cleaved InGaAsP lasers on an InP substrate, etchants of $4\text{H}_2\text{SO}_4:\text{1H}_2\text{O}_2:\text{1H}_2\text{O}$ and concentrated HCl were used to selectively etch layers of InGaAsP and InP, respectively. Stripe-geometry, buried-heterostructure lasers fabricated by this method to emit at $1.3\text{ }\mu\text{m}$ had threshold currents as low as 20 mA and differential quantum efficiencies as high as 60% [44]. The yield in the cleaving process step was 77% .

There is usually no difficulty in making electrical contact to the p and n sides of the junction of a discrete laser diode, but when the diode is monolithically integrated into an optical integrated circuit it may not be possible to obtain a return path for current through the substrate (for example, when a semi insulating substrate is used). In such cases a transverse-junction-stripe (TJS) laser is called for. The TJS laser [45–47], shown in Fig. 14.14, features a transverse junction which has both the p and the n electrical contacts on the top surface. A double heterostructure, stripe-geometry laser cavity is produced in n-type layers grown on a semi-insulating substrate.

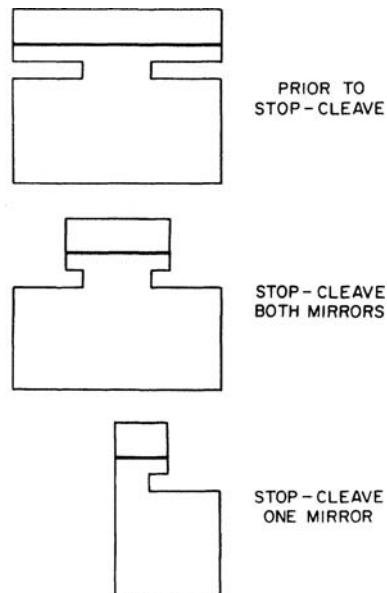


Fig. 14.13 Laser diodes with stopcleaved facets [43]

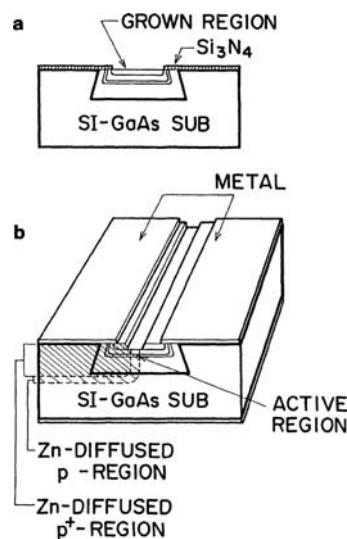
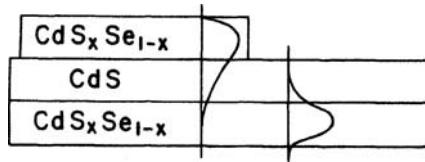


Fig. 14.14 A transverse junction stripe (TJS) laser [45]

Fig. 14.15 A monolithic laser/waveguide structure in CdSeS [49]



Then masked Zn diffusion is used to produce a lateral p-n junction. GaAlAs TJS lasers fabricated by MOCVD growth have exhibited threshold current densities as low as 25 mA and external differential quantum efficiencies of 40% [46]. A TJS laser has been produced in GaInAsP on a semi-insulating InP substrate, by Oe et al. [48]. Threshold current of 10 mA and maximum CW output power of 10 mW were achieved. Because of the surface configuration of the electrodes TJS lasers are particularly suitable for use at microwave modulation frequencies, with the modulation drive signal being fed to the laser over a metallic strip line deposited directly onto the surface of a semi-insulating GaAs or InP substrate.

A different approach to laser/waveguide integration was demonstrated by Kawabe et al. [49] An optically pumped $\text{CdS}_x\text{Se}_{(1-x)}$ laser was coupled to a waveguide of the same material by optical tunneling through a CdS confining layer, as shown in Fig. 14.15. Stripe geometry optical pumping by a nitrogen laser above a threshold of 70 kW/cm^2 produced lasing in the $\text{CdS}_x\text{Se}_{(1-x)}$. A single transverse optical mode was observed in the laser, and this mode was coupled into the waveguide as shown in Fig. 14.15 for confining (CdS) layer thickness of $2.5 - 6 \mu\text{m}$. This type of coupling, which can also be used for coupling diode lasers to a waveguide, is called evanescent wave coupling.

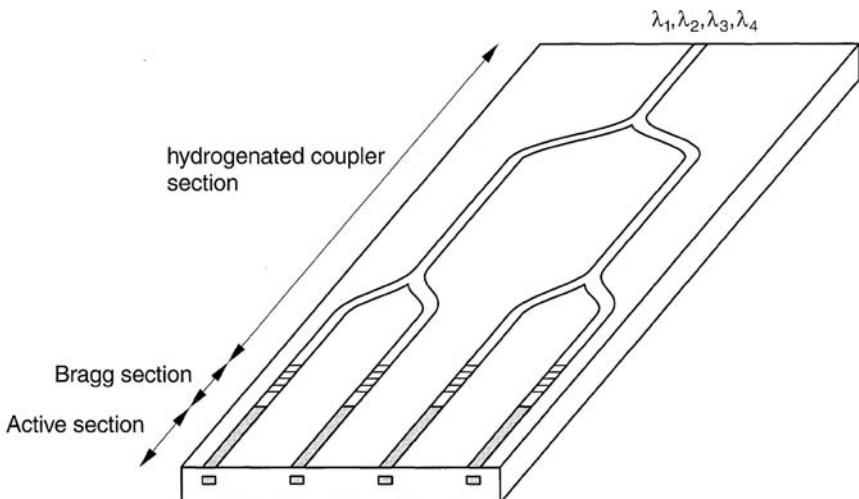


Fig. 14.16 Four-wavelength selectable OIC source [51] © 1999 IEEE

The difficulty of providing Fabry-Perot type feedback can be entirely avoided by employing distributed feedback from a Bragg diffraction grating of the type described in Chapter 10. For example, Aiki et al. [50] have used that approach to produce a frequency multiplexing light source incorporating six monolithically integrated distributed feedback (DFB) lasers coupled to GaAlAs waveguides, and Talneav et al. [51] have made an OIC in GaInAsP by MOVPE that features 4 lasers coupled to a single output by branching waveguides, as shown in Fig. 14.16. In addition to facilitating monolithic coupling, DFB lasers also provide the advantages of narrower linewidth and better mode stability, as compared to lasers with Fabry-Perot cavities. Because of their importance in optical integrated circuits, DFB lasers are described in detail in Chapter 15.

14.5 Reliability

A key characteristic of laser diodes, which strongly affects system design, is the reliability of the device. Laser diodes are very highly stressed in terms of the current densities and optical field strengths which exist within them, and they generally exhibit some gradual degradation of their operating characteristics as they age, as well as being subject to catastrophic failure under certain conditions.

14.5.1 Catastrophic Failure

Laser diodes are subject to catastrophic failure that occurs when the optical power density at the mirror facets exceeds a certain critical level, generally accepted to be in the range of 2 to 3×10^6 W/cm² for GaAlAs lasers and slightly higher for GaInAsP devices. Mechanical damage, in the form of pits and protrusions, usually occurs first near the center of the facet, where the optical field strength is maximum. The threshold power density above which catastrophic failure occurs varies for different lasers, due in part to the fact that the presence of grown-in initial defects in the facet surface enhances the formation of further damage. Catastrophic damage can be reduced by coating the laser facets with a dielectric such as SiO₂, Si₃N₄ or Al₂O₃ is closely matched to GaAs in its coefficient of thermal expansion and has high thermal conductivity. Operating the laser on a pulsed basis rather than CW also results in reduced catastrophic damage.

Catastrophic damage can, of course, be avoided by merely operating the laser well below the damage threshold. However, it must be remembered that transients of only a few microseconds duration are sufficient to destroy the laser. Hence, turn-on transients must be filtered out in cw operation, and ringing must be avoided in pulsed applications. In addition, power supply filtering must be adequate to absorb any random transients that can occur on the line.

14.5.2 Gradual Degradation

The problem of gradual degradation, in which the threshold current density of the laser increases and differential quantum efficiency decreases as the device is operated, was so severe in early laser diodes that device lifetime was only 1 to 100 h, and it was therefore believed that a practical, reliable device would never be made. However, methodical work in a number of different laboratories over a period of years has led to the identification and elimination of most degradation mechanisms, so that lasers with projected lifetimes of 10^6 h are now being made [52]. The gradual degradation of laser diodes results from the formation of defects within the active region that act as non-radiative recombination centers [53]. The failure modes and mechanisms in laser diodes have been reviewed by Fukuda [54].

14.6 Vertical Cavity Lasers

Vertical cavity lasers are a relatively new type of semiconductor laser that emit light through their surface rather than through their endfaces. In fact, they are often called vertical-cavity surface-emitting lasers (VCSELs). The typical device structure of a VCSEL is shown in Fig. 14.17. The key elements are an active layer, two DFB mirrors and a contact window to allow light to be emitted from the surface. In the particular device shown, a strain-compensated, quantum-well, 7-layer active region is sandwiched between two AlGaAs/GaAs mirrors by wafer fusion [55].

Typically, the active region is a multilayer, double heterostructure diode, as in a conventional Fabry-Perot endface laser. However, in this case, lightwaves travel in the direction perpendicular to the junction plane. They are reflected by top and bottom “mirrors” which consists of multilayer structures, alternating layers of materials with differing indices of refraction that are approximately one half wavelength thick

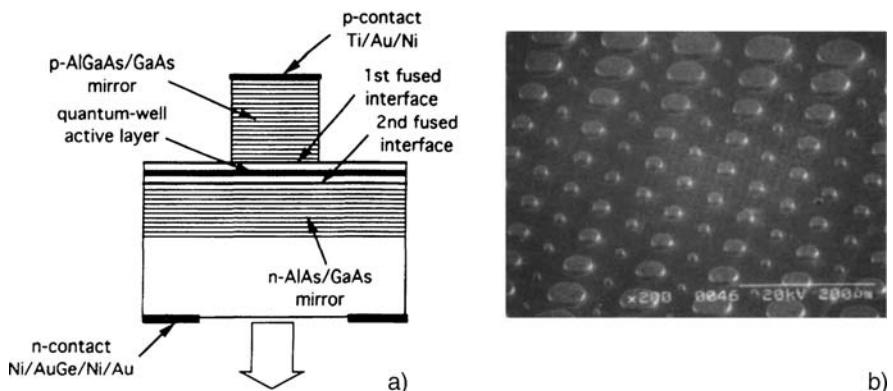


Fig. 14.17 a,b Vertical cavity laser **a** cross-sectional view **b** surface view of array [55] © 1995 IEEE

(at the lasing wavelength in the material). For this spacing, the lightwaves reflected from each interface between layers positively reinforce waves reflected from all of the others in the reverse direction, so an effective “mirror” is formed. The overall reflectivity of the multilayer structure depends on the reflectivity at each interface and on the number of layers. This type of distributed feedback (DFB) reflector will be described in detail in Chapter 15, which is devoted entirely to DFB lasers. The reflecting semiconductor layers are usually grown by either MOCVD or MBE, since sub-micron thickness is required. In Fig. 14.17 typical layers of GaAs and GaAlAs are shown but other materials such as GaInAsP/InP are often used. In some cases the reflecting layers can be thin films of dielectrics.

VCSELs have a number of advantages over Fabry-Perot endface diode lasers in certain applications. They can be made in a circular shape, as shown in Fig. 14.17b, with a diameter matched to the core diameter of an optical fiber to optimize coupling. Because of their relatively largely emitting area, the beam remains relatively well collimated and does not diverge due to diffraction as does that of an endface-emitting laser diode. Also, VCSELs can be conveniently arranged into a surface array to couple to a fiber ribbon or bundle, or to provide direct chip-to-chip coupling through the air. The surface orientation of emission also facilitates integration of a VCSEL with planar electronic circuitry. For example, Krishnamoorthy et al. [56] have used flip-chip bonding to integrate VCSELs with conventional silicon CMOS circuits as shown in Fig. 14.18.

The integration of VCSELs and CMOS circuitry provides an effective way to make IOC transceivers in which low-area, low-power, high-speed CMOS circuits

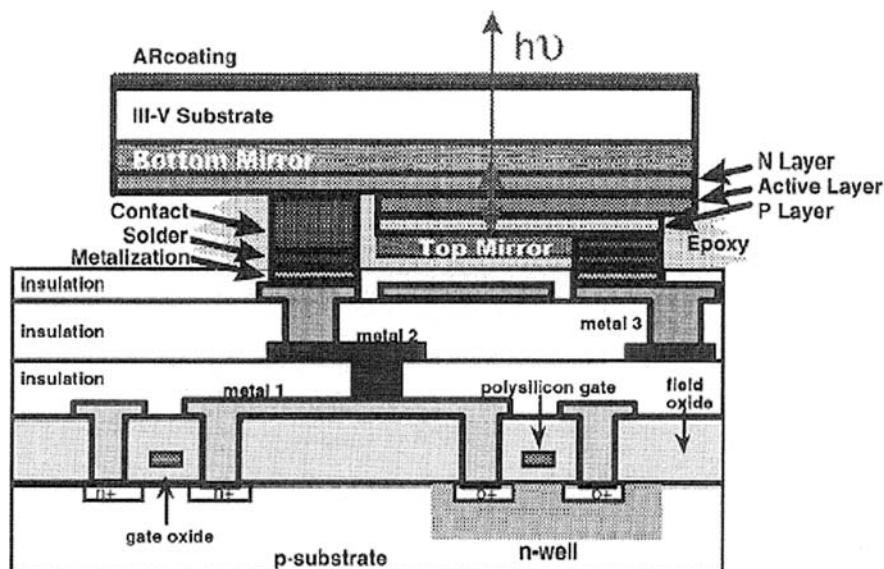


Fig. 14.18 Cross-sectional view of VCSEL/CMOS integration [56] © 1999 IEEE

are used to modulate VCSELs in a surface-emitting array. The circuit of Fig. 14.18 operated at 1.25 Gb/s with a bit error rate (BER) less than 10^{-11} .

VCSELs are part of the general class of lasers known as distributed Bragg reflector (DBR) lasers because their required positive feedback is provided by distributed Bragg-type reflections. The phenomenon of distributed feedback (DFB) and its use in DBR lasers are explained in detail in the next chapter.

Problems

- 14.1 Sketch the cross sectional view of a double-heterostructure GaAlAs laser diode. Choose the thickness and aluminum concentration in each layer required to produce light emission at $\lambda_0 = 8500 \text{ \AA}$, along with field confinement to the active layer for the lowest order TE and TM modes only.
- 14.2 Compare the operating characteristics of homojunction, SH and DH lasers, describing the advantages and disadvantages of each type.
- 14.3 Derive an expression for the threshold current density J_{th} in a heterojunction laser that has an active layer thickness d , light emitting layer thickness D , average loss coefficient α , and reflecting endfaces with different reflectivities R_1 and R_2 . Use the symbols defined in the text for the other parameters, i.e., η_q , λ_0 , $\Delta\nu$, etc.
- 14.4 For small active layer thickness d , the confinement factor d/D may be written as

$$d/D \cong \gamma d.$$

The compositional dependence of the refractive index of $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ at $\lambda_0 = 0.90 \mu\text{m}$ can be represented by

$$n = 3.590 - 0.710x + 0.091x^2.$$

- (a) Develop an expression for the confinement factor as a function of x , d and λ_0 for a three-layer, symmetric $\text{GaAs}-\text{Ga}_{(1-x)}\text{Al}_x\text{As}$ DH laser with $n_a = 3.590$.
- (b) For a DH laser with $d = 0.1 \mu\text{m}$, $x = 0.1$ and $\lambda_0 = 0.90 \mu\text{m}$, what is the magnitude of the confinement factor?
- (c) Compare your answer to Part b, with the data given in Fig. 14.4
- 14.5 Sketch the energy-band diagram for the DH laser of Fig. 14.5 when the laser is forward biased to above threshold and show the carrier confinement.
- 14.6 What are the advantages of a stripe-geometry laser as compared to a broad-area DH laser?

- 14.7 The external *differential quantum* efficiency of a GaAs laser diode emitting at a wavelength of $0.9 \mu\text{m}$ is 30%. The applied voltage is 2.5 volts. Calculate the external *power* efficiency of the device.
- 14.8 A certain semiconductor p-n junction laser when biased as required has a gain coefficient $g = 70 \text{ cm}^{-1}$, a loss coefficient $\alpha = 5 \text{ cm}^{-1}$, and a length between Fabry-Perot endfaces of $L = 300 \mu\text{m}$. It is desired to add antireflection coatings to the endfaces. What is the smallest endface reflectivity R that can be used without going below the lasing threshold?
 (Assume that both endfaces have the same reflectivity, and that the confinement factor $d/D = 1$.)
- 14.9 (a) List the advantages of a double-heterostructure laser as compared to a p-n junction laser diode without heterojunctions.
 (b) Explain the reasons for the features that you listed in part (a).
- 14.10 Explain why the double-heterostructure semiconductor laser is considered to be so important to modern-day optical communications systems that its inventors were awarded the Nobel Prize in Physics in 2000.

References

1. G. Diemer, B. Böger: Physics **29**, 600 (1963)
2. T. Pecany: Phys. Stat. Sol. **6**, 651 (1964)
3. H. Kroemer: IEEE Proc. **51**, 1782 (1963)
4. Zh.I. Alferov: Sov. Phys.-Solid State **7**, 1919 (1966)
5. I. Hayashi, M.B. Panish, P. Foy: IEEE J. **QE-5**, 211 (1969)
6. H. Kressel, H. Nelson: RCA Rev. **30**, 106 (1969)
7. Zh.I. Alferov, V. Andreev, E. Portnoi, M. Trukhan: Sov. Phys. – Semicond. **3**, 1107 (1969)
8. K. Sheger, A. Milnes, D. Feucht: Proc. Int'l Conf. on Chem. Semicond. Hetero-junction Layer Structures, Budapest (Hung. Acad. Sci., Budapest 1970) Vol. 1, p. 73
9. P.H. Holloway, T.J. Anderson: *Compound Semiconductors: Growth, Processing and Devices* (CRC Press, Boca Raton, FL 1989) p. 115
10. Q.H.F. Vrehen: J. Phys. Chem. Solids **29**, 129 (1968)
11. H. Yonezu, I. Sakuma, Y. Nannich: Jpn. J. Appl. Phys. **9**, 231 (1970)
12. A. Yariv, R.C.C. Leite: Appl. Phys. Lett. **2**, 173 (1963)
13. H.C. Casey Jr., M.B. Panish: *Heterostructure Lasers*, Pt. B: Materials and Operating Characterizations (Academic, New York 1978) pp. 109–132
14. H. Kroemer: IEEE Trans. **ED-39**, 2635 (1992)
15. K. Iga, S. Kinoshita: *Semiconductor Lasers and Related Epitaxies*, Springer Ser. Mater. Sci., Vol. 30 (Springer, Berlin, Heidelberg 1995)
16. A. McWhorter: Solid State Electron. **6**, 417 (1963)
17. H.C. Casey Jr., M.B. Panish: *Heterostructure Lasers*, Pt. A: Fundamental Principles (Academic, New York 1978) pp. 54–57
18. J. Kongas, P. Savolainen, M. Toivonen, S. Orsila, P. Corvini, M. Jansen, R. Nabiev, M. Pesa: High-efficiency AlGaInP single-mode laser. IEEE Photonics Tech. Lett. **10**, 1533 (1998)
19. R.J. Lang, N.W. Carlson, E. Beyer, M. Obara: Introduction to the issue on high-power and high brightness lasers. IEEE J. Select. Topics Quant. Electron. **6**, 561 (2000)
20. W. Schulz, R. Poprawe: Manufacturing with novel high-power diode lasers. IEEE J. Select. Topics Quant. Electron. **6**, 696 (2000)
21. D. Greenaway, G. Harbecke: *Optical Properties and Band Structure of Semiconductors* (Pergamon, Oxford 1968) p. 67

22. D.N. Payne, W.A. Gambling: Electron. Lett. **11**, 176 (1975)
23. M.J. Li, C. Saravacos: Optical fiber design for field-mountable connectors. IEEE J. Lightwave Tech. **18**, 314 (2000)
24. H. Kressel (ed.): *Semiconductor Devices for Optical Communications*, 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982) pp. 285–289
25. H. Mani, A. Joullie, G. Boissier, E. Tournie, F. Pitard, C.A. Ailbert: Electron. Lett. **24**, 1542 (1988)
26. R.V. Martinelli: LEOS'88, Santa Clara, CA. Digest p. 55
27. H. Ebe, Y. Nishijima, K. Shinohara: 11th IEEE Int'l Conf. on Semicond, Lasers, Boston, MA (1988) Digest p. 68
28. I.T. Sorokina, K.L. Vodopyanov, (eds.): *Solid-State Mid-Infrared Laser Sources*, Topics in Applied Physics Series, vol. 89 (Springer, Berlin, Heidelberg, 2003)
29. J.C. Dyment: Appl. Phys. Lett. **10**, 84 (1967)
30. L.A. D'Asaro: J. Lumin. **7**, 310 (1973)
31. H. Yonezu, I. Sakuma, K. Kobayashi, T. Kamejima, M. Ueno, Y. Nannicki: Jpn. J. Appl. Phys. **12**, 1585 (1973)
32. T. Tsukada: J. Appl. Phys. **45**, 4899 (1974)
33. M. Nakamura: IEEE Trans. **CAS-26**, 1055 (1979)
34. N. Chinone: J. Appl. Phys. **48**, 3237 (1977)
35. K. Seki, T. Kamiya, H. Yanai: Trans. IECE (Jpn.) E-**62**, 73 (1979)
36. W.O. Schlosser: Bell. Syst. Tech. J. **52**, 887 (1973)
37. W.T. Tsang, R.A. Logan, M. Illegems: Appl. Phys. Lett. **32**, 311 (1978)
38. T. Kobayashi, H. Kawaguchi, Y. Furukawa: Jpn. J. Appl. Phys. **16**, 601 (1977)
39. I.P. Kaninow, R.S. Tucker: Mode-controlled semiconductor lasers, in *Guided-Wave Optoelectronics*, T. Tamir, 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) pp. 211–263
40. R. Baets: Solid State Electron. **30**, 1175 (1987)
41. C.E. Hurwitz, J.A. Rossi, J.J. Hsieh, C.M. Wolfe: Appl. Phys. Lett. **27**, 241 (1975)
42. L.A. Koszi, A.K. Chin, B.P. Segner, T.M. Shen, N.K. Dutta: Electron. Lett. **21**, 1209 (1985)
43. A. Antreasyn, C.Y. Chen, R.A. Logan: Electron. Lett. **21**, 405 (1985)
44. A. Antreasyn, S.G. Napholtz, D.P. Wilt, P.A. Garbinski: IEEE J. **QE-22**, 1064 (1986)
45. M. Ishii, K. Kamon, M. Shimazu, M. Mihara, H. Kumabe, K. Isshiki: Electron. Lett. **23**, 179 (1987)
46. M. Ishii, K. Kamon, M. Shimazu, M. Mihara, H. Kumabe, K. Isshiki: Optoelectron. – Devices Technol. **2**, 83 (1987)
47. S. Lathi, K. Tanaka, T. Morita, S. Inoue, H. Kan, Y. Yamamoto: Transverse-junction-stripe GaAs-AlGaAs lasers for squeezed light generation. IEEE J. Quant. Electron. **35**, 387 (1999)
48. K. Oe, Y. Noguchi, C. Canea: IEEE Photon. Tech. Lett. **6**, 479 (1994)
49. M. Kawabe, H. Kotani, K. Masuda, S. Namba: Appl. Phys. Lett. **26**, 46 (1975)
50. K. Aiki, M. Nakamura, J. Umeda: Appl. Phys. Lett. **29**, 506 (1976)
51. A. Talneau, M. Allovon, N. Bouadma, S. Slempkes, A. Ougazzaden, H. Nakajima: Agile and fast switching monolithically integrated four wavelength selectable source at $1.55\text{ }\mu\text{m}$. IEEE Photonics Tech. Lett. **11**, 12 (1999)
52. M. Krakowski, R. Blondeau, J. Ricciardi, J. Hirtz, M. Razeghi, B. de Cremoux: OSA/IEEE OFC/IGWO'86, Atlanta, GA. Paper TU33
53. D.I. Babic, K. Streubel, R.P. Mirin, N.M. Margalit, J.E. Bowers, E.L. Hu, D.E. Mars, L. Yang, K. Carey: Room-temperature continuous-wave operation of $1.54\text{ }\mu\text{m}$ vertical-cavity lasers. IEEE Photonics Tech. Lett. **7**, 1225 (1995)
54. M. Fukuda: Historical overview and future of optoelectronics reliability for optical fiber communications systems, Microelectron. Reliability **40**, 27 (2000)
55. T. Kallstenius, A. Landstedt, U. Smith, P. Granstrand: Role of nonradiative recombination in the degradation of InGaAsP/InP-based bulk lasers. IEE J. Quant. Electr. **36**, 1312 (2000)

56. A.V. Krishnamoorthy, L.M.F. Chirovsky, W.S. Hobson, R.E. Leibenguth, B.P. Hui, G.J. Zyzdzik, K.W. Goossen, J.D. Wynn, B.J. Tseng, J. Lopata, J.A. Walker, J.E. Cunningham, L.A. D'Asaro: Vertical-cavity surface-emitting lasers flip-chip bounded to gigabit-per-second CMOS circuits. *IEEE Photonics Tech. Lett.* **11**, 128 (1999)

Supplementary Reading on Heterojunction Lasers

- Z.I. Alferov: Nobel Lecture: The double heterostructure concept and its applications in physics, electronics and technology, *Rev. Modern Phys.* **73**, 767–782 (2001)
- P. Bhattacharya: *Semiconductor Optoelectronic Devices*, 2nd edn. (Prentice Hall, Upper Saddle River, New Jersey 1997) Chap. 7
- J.K. Butler (ed.): *Semiconductor Injection Lasers* (IEEE Press, New York 1980)
- H.C. Casey Jr., M.B. Panish: *Heterostructure Lasers* (Academic, New York 1978)
- N. Grote, H. Venghaus (eds.): *Fiber Optic Communication Devices*, Springer Series in Photonics, vol. 4 (Springer, Berlin, Heidelberg, 2001)
- H. Kressel (ed.): *Semiconductor Devices for Optical Communication*, 2nd edn., Topics Appl. Phys., Vol. **39** (Springer, Berlin, Heidelberg 1982) Chap. 2
- H. Kressel, J.K. Butler: *Semiconductor Lasers and Heterojunction LEDs* (Academic, New York 1977)
- W.B. Leigh: *Devices for Optoelectronics* (Marcel Dekker, New York 1996) Chap. 3
- T. Tamir (ed.): *Guided-Wave Optoelectronics*, 2nd edn., Springer Ser. Electron. Photon., Vol. 26 (Springer, Berlin, Heidelberg 1990) Chap. 5
- A. Yariv: *Optical Electronics*, 4th edn. (Saunders College Publishing-HRW, Philadelphia 1991) Chap. 15

Chapter 15

Distributed-Feedback Lasers

Most of the lasers that have been described so far depend on optical feedback from a pair of reflecting surfaces, which form a Fabry-Perot etalon. In an optical integrated circuit, in which the laser diodes are monolithically integrated within the semiconductor wafer, it is usually very difficult to form such reflecting surfaces. They can be formed by etching or cleaving, as described in Chapter 14. However, the planar surface of the wafer is then disrupted, which leads to difficulties in fabricating electrical connections and heat sinks. An alternative approach, which utilizes distributed-feedback (DFB) from a Bragg-type diffraction grating, provides a number of advantages while still utilizing a planar surface geometry.

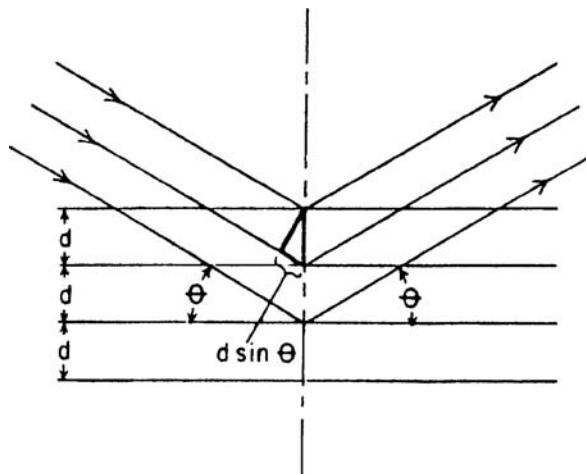
15.1 Theoretical Considerations

The use of a Bragg-type diffraction grating to deflect an optical beam in a modulator is described in Chapter 10. In that case, the grating structure is usually produced by inducing a periodic change in the index of refraction by means of either the electro-optic or acousto-optic effect. In DFB lasers, the grating is usually produced by corrugating the interface between two of the semiconductor layers that comprise the laser. This corrugation provides 180° reflection at certain specific wavelengths, depending on the grating spacing.

15.1.1 Wavelength Dependence of Bragg Reflections

The basis for selective reflection of certain wavelengths can be understood by referring to Fig. 15.1 which illustrates the reflection of an incident plane wave by a series of reflectors spaced at a distance d . In the original case of x-ray diffraction that was considered by Bragg [1], these reflectors were the atomic planes of a crystalline lattice. However, the same effect is observed in the case of reflection from a corrugated grating formed in the junction plane of a semiconductor laser as shown in Fig. 15.2. It is obvious from the diagram shown in Fig. 15.1 that, in order to maintain the phase coherence of the plane wavefront (normal to the rays shown) and

Fig. 15.1 Diagram of Bragg reflection from a periodic structure



thereby avoid destructive interference, the path lengths for reflections from successive reflectors must differ by an integral number of full wavelengths. Thus, from geometrical considerations we obtain the Bragg relation [1], given by

$$2d \sin \theta = l\lambda, \quad l = 1, 2, 3, \dots, \quad (15.1)$$

where θ is the angle formed by the incident ray and the reflector, and λ is the optical wavelength in the medium. To adapt the relation (15.1) to the case of 180° reflection by a grating in a DFB laser, it is only necessary to let d equal the grating spacing Λ , let λ equal λ_0/n_g , where n_g is the effective index in the waveguide for the mode under consideration, and let θ equal 90° . Under these assumptions (15.1) becomes

$$2\Lambda = l(\lambda_0/n_g), \quad l = 1, 2, 3, \dots \quad (15.2)$$

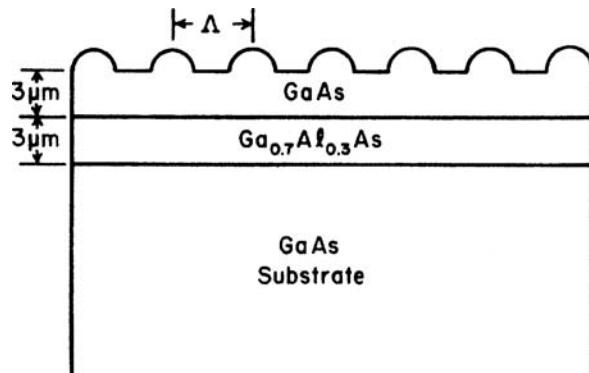


Fig. 15.2 Cross-sectional view of an optically pumped DFB laser

The vacuum wavelength of light that will be reflected through 180° by such a grating is therefore

$$\lambda_0 = \frac{2\Lambda n_g}{l}, \quad l = 1, 2, 3, \dots, \quad (15.3)$$

Although the grating is capable of reflecting many different longitudinal modes, corresponding to the various values of l , usually only one mode will lie within the gain bandwidth of the laser. In fact, because of the difficulty of fabricating a first order ($l = 1$) grating, usually a third-order grating is used, as is discussed in detail in Section 15.2.

15.1.2 Coupling Efficiency

The fraction of optical power that is reflected by a grating such as that shown in Fig. 15.2 depends on many factors, including the thickness of the waveguiding layer, the depth of the grating teeth and the length of the grating region. The determination of the fraction of reflected optical power is mathematically quite complex. It is best approached by using the coupled-mode theory, as has been done by Kogelnik and Shank [2, 3], and by Yariv [4], although the Bloch wave formalism has also been used [5]. In the coupled-mode analysis, it is assumed that the grating is only weakly coupled to the optical modes, i.e., the basic mode shape is only slightly perturbed. Using this approach, Yariv [4] has shown that for a rectangular grating cross section of depth a , as shown in Fig. 15.3, the coupling can be characterized by a coupling coefficient κ , given by

$$\kappa = \frac{2\pi^2}{3l\lambda_0} \frac{(n_2^2 - n_1^2)}{n_2} \left(\frac{a}{t_g} \right)^3 \left[1 + \frac{3(\lambda_0/a)}{2\pi(n_2^2 - n_1^2)^{1/2}} + \frac{3(\lambda_0/a)^2}{4\pi^2(n_2^2 - n_1^2)} \right], \quad (15.4)$$

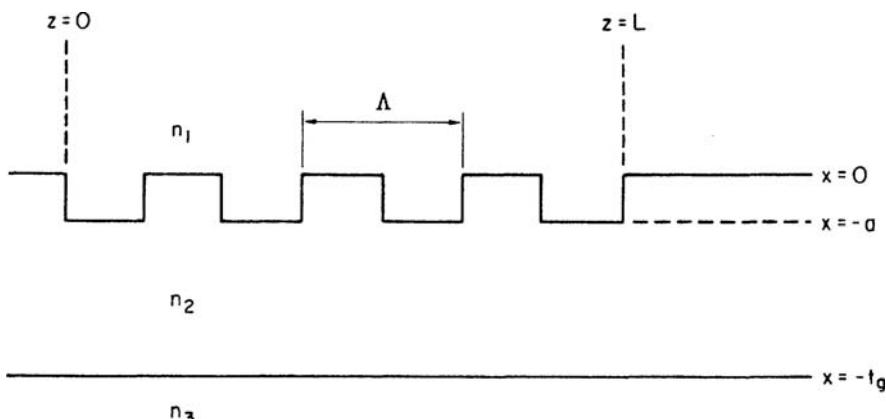


Fig. 15.3 An optical waveguide with a rectangular grating for distributed feedback [4]

where t_g is the thickness of the waveguiding layer with index n_2 , and l is the order of the harmonic responsible for the coupling, given by

$$l \cong \frac{\beta_m \Lambda}{\pi}, \quad (15.5)$$

where β_m is the phase constant of the particular mode under consideration. The expressions (15.4) and (15.5) are based on the assumption that β_m is approximately equal to $l\pi/\Lambda$, i.e. the Bragg condition (15.3) is satisfied, so that Bragg reflection occurs between the forward and backward traveling modes of order m .

For the first order ($l = 1$) case, the reflection is essentially limited to coupling between only the forward and backward waves traveling in the $\pm z$ direction. Then the amplitude of the incident wave A_m^+ and that of the reflected wave A_m^- are functions of the distance z that the incident wave travels under the grating, given by the expressions [4]

$$A_m^+(z) = A_m^+(0) \frac{\cosh[\kappa(z - L)]}{\cosh(\kappa L)}, \quad (15.6)$$

and

$$A_m^-(z) = A_m^+(0) \frac{\kappa}{|\kappa|} \frac{\sinh[\kappa(z - L)]}{\cosh(\kappa L)}, \quad (15.7)$$

where L is the length of the grating. The incident and reflected optical powers are, of course, given by $|A_m^+(z)|^2$ and $|A_m^-(z)|^2$, respectively. For sufficiently large arguments of the hyperbolic functions in (15.6) and (15.7), the incident optical power decreases exponentially with z , as power is reflected into the backward traveling wave. The effective reflection coefficient R_{eff} and corresponding transmissions T_{eff} at the Bragg wavelength are given by

$$R_{\text{eff}} = \left| \frac{A_m^-(O)}{A_m^+(O)} \right|^2, \quad (15.8)$$

and

$$T_{\text{eff}} = \left| \frac{A_m^+(L)}{A_m^+(O)} \right|^2. \quad (15.9)$$

The relations (15.6), (15.7), (15.8) and (15.9), strictly speaking, are applicable only in the case of first-order Bragg reflection. For higher-order gratings, with periodicity Λ given by

$$\Lambda = \frac{l\lambda_0}{2n_g}, \quad l > 0, \quad (15.10)$$

some optical energy is coupled into waves traveling in directions out of the waveguide plane, as shown in Fig. 15.4. These waves represent a power loss from the lasing mode, and hence produce reductions of both R_{eff} and T_{eff} .

Scifres et al. [6] have analyzed the reflections from higher-order gratings as shown in Fig. 15.5. Consider a wave traveling to the right in the corrugated waveguide. In order for the rays scattered from successive teeth to all be in phase, the path lengths must all be integral multiples of the wavelength. Hence a conditions is imposed which requires that

$$b + \Lambda = \frac{l' \lambda_0}{n_g}, \quad l' = 1, 2, \dots, \quad (15.11)$$

where b is the distance shown in Fig. 15.5.

From geometrical considerations, it can be seen that

$$b = \Lambda \sin \theta, \quad (15.12)$$

where θ is the angle that the scattered wavefront makes with the plane of the waveguide. Thus, from (15.11) and (15.12), θ is given by

$$\sin \theta = \frac{l' \lambda_0}{n_g \Lambda} - 1, \quad l' = 0, 1, 2, \dots, l. \quad (15.13)$$

Fig. 15.4 a-d Reflected wave directions for higher-order Bragg diffraction gratings

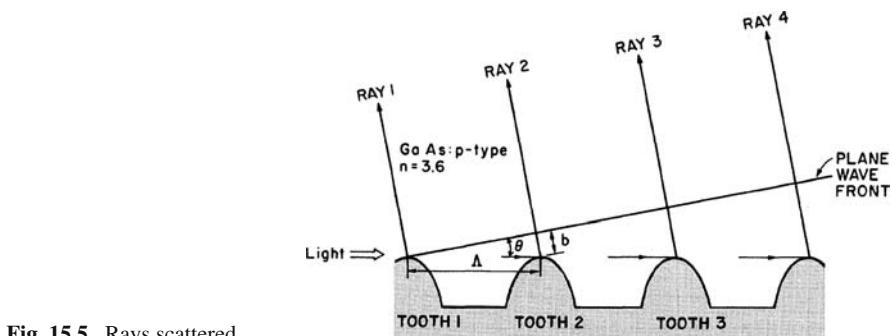
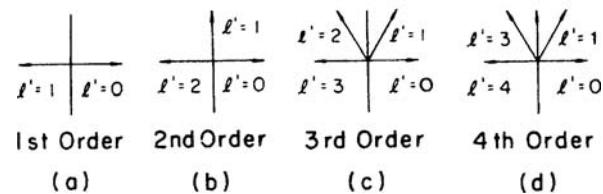


Fig. 15.5 Rays scattered from a waveguide with a higher-order Bragg grating [6]

Consider the case of second-order Bragg diffraction, for which Λ equals λ_0/n_g . In this case (15.13) becomes

$$\sin \theta = l' - 1, l' = 0, 1, 2. \quad (15.14)$$

The solution for $l' = 0$ corresponds to light scattered in the forward direction, while that for $l' = 2$ gives backward scattering useful for DFB. However, the $l' = 1$ solution yields $\theta = 0$, corresponding to light emitted normal to the waveguide plane. Similar solutions can be obtained for higher-order Bragg reflection gratings. The coupling coefficients have been given by Streifer et al. [7] for a number of typical cases of scattering from higher-order Bragg gratings.

Generally, the strong coupling in the transverse direction produced by second-order Bragg gratings is undesirable in a DFB laser. Hence, a first-order grating is required to yield optimum performance. If fabrication tolerances prevent the use of a first-order grating, a third-order grating should be used, since coupling to the out-of-plane beams is much weaker than in the second-order case. One exception to this rule is a special, low-divergence-angle laser, demonstrated first by Scifres et al. [6]. They intentionally used a fourth-order grating to produce a strong transverse coupling in a GaAlAs/GaAs SH laser. The Fabry-Perot end faces of the laser were coated to make them 98% reflective, and the laser output was taken in the transverse direction. Because of the relatively large size of the light emitting surface, diffraction spreading was minimized, and a beam with angular divergence of only 0.35° was produced.

15.1.3 Lasing with Distributed Feedback

It was shown in the previous section that a corrugated grating structure can provide 180° Bragg reflection of an optical wave traveling in close proximity to the grating. If the medium in which the wave is traveling has optical gain, such as that of the inverted population region of a laser diode, this distributed feedback can result in lasing. In a medium with an exponential gain constant γ , the amplitudes of the incident and reflected waves are given by the expressions [8]

$$E_i(z) = E_0 \frac{\{(\gamma - i\Delta\beta) \sinh[S(L-z)] - S \cosh[S(L-z)]\} e^{i\beta_0 z}}{(\gamma - i\Delta\beta) \sinh(SL) - S \cosh(SL)}, \quad (15.15)$$

and

$$E_r(z) = E_0 \frac{\kappa e^{i\beta_0 z} \sinh[S(L-z)]}{(\gamma - i\Delta\beta) \sinh(SL) - S \cosh(SL)}, \quad (15.16)$$

where

$$S^2 \equiv |\kappa|^2 + (\gamma - i\Delta\beta)^2. \quad (15.17)$$

The parameter E_0 is the amplitude of a single mode (the one for which net gain is greatest) incident on the grating of length L at $z = 0$, while $\Delta\beta$ is given by

$$\Delta\beta = \beta - \beta_0, \quad (15.18)$$

where β_0 is the phase constant at the Bragg wavelength.

The oscillation condition for the DFB laser [2, 3] corresponds to the case for which both the transmittance $E_i(L)/E_i(0)$ and the reflectance $E_r(0)/E_i(0)$ become infinite. From (15.15), (15.16) and (15.17), it can be shown that this condition is satisfied when

$$(\gamma - i\Delta\beta) \sinh(SL) = S \cosh(SL). \quad (15.19)$$

In general, (15.19) can be solved to determine the threshold values of $\Delta\beta$ and γ only by numerical solution [2, 3]. However, in the special case of high gain ($\gamma \gg |\kappa|$, $\Delta\beta$) it can be shown [8, p. 381] that the oscillating mode frequencies are given by

$$(\Delta\beta_m)L \cong -(m + \frac{1}{2})\pi. \quad (15.20)$$

Since it is known from basic definitions that

$$\Delta\beta \equiv \beta - \beta_0 \cong \frac{(\omega - \omega_0)n_g}{c}, \quad (15.21)$$

(15.20) can be written in the form

$$\omega_m = \omega_0 - (m + \frac{1}{2})\frac{\pi c}{n_g L}, \quad (15.22)$$

where ω_0 is the Bragg frequency, and $m = 0, 1, 2, \dots$. It is interesting to note in (15.22) that no oscillation can occur at exactly the Bragg frequency. The mode frequency spacing is given by

$$\omega_{m-1} - \omega_m \cong \frac{\pi c}{n_g L}. \quad (15.23)$$

15.2 Fabrication Techniques

The grating structure of a DFB laser is usually formed by masking and then etching the waveguide surface, either chemically or by ion-beam sputtering. Sometimes the two methods are combined in a single process called chemically assisted ion beam etching (CAIBE) [9]. The process is generally the same as that described in

Section 7.4.2 for grating-coupler fabrication. However, some additional factors must be considered in order to produce an efficient laser.

15.2.1 Effects of Lattice Damage

In an integrated DFB laser, such as the one shown in Fig. 15.6, it is necessary to grow one or more epitaxial layers on top of the waveguide layer after the grating has been formed. Lattice damage created during the grating fabrication process will, of course, be detrimental to the optical quality of the subsequently grown epitaxial layers, because point defects and dislocations tend to propagate from the grating region into the epi layer. These defects act as centers for optical absorption and non-radiative recombination. Thus, they act to reduce quantum efficiency and increase threshold current density. The deleterious effects of defects generated during grating fabrication present somewhat of a dilemma to the integrated circuit designer when choosing the fabrication technique, because ion-beam sputter-etching produces the most uniform and controllable grating profile, but chemical etching yields gratings with much less lattice damage.

15.2.2 Grating Location

The approach that is most often used to avoid the effects of lattice damage is to physically separate the grating from the active layer of the laser. For example, Scifres et al. [10] made the first p-n junction injection DFB laser, by using a diffusion process that separated the p-n junction from the grating as shown in Fig. 15.7. The n-GaAs substrate was first corrugated by ion-beam sputter-etching to produce a third order grating with 3500 Å periodicity. Then the p-GaAlAs layer was grown over the substrate. Zinc was used as the p-type dopant in the GaAlAs because of its anomalously fast diffusion in GaAs. Hence, during the growth of the GaAlAs layer, the zinc diffused into the substrate to produce a p-n junction at about 1 μm below

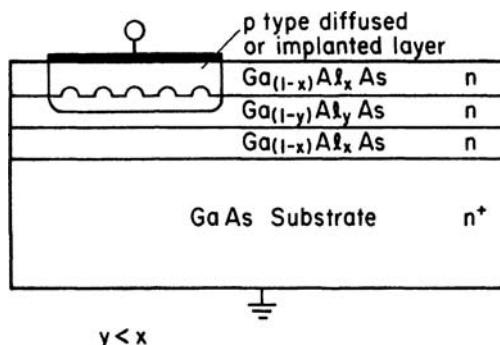
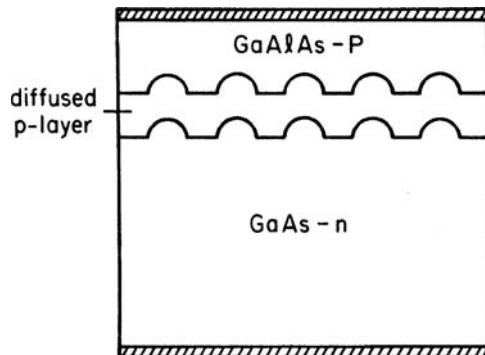


Fig. 15.6 Integrated DFB laser

Fig. 15.7 Diffused junction DFB laser [10]



the GaAs-GaAlAs (corrugated) interface. The resulting single heterostructure DFB lasers exhibited threshold densities that were typically equal to those of similar SH lasers with cleaved end faces. Thus, one can conclude that the effects of defects generated during grating fabrication were not severe.

Another approach to separating the grating from the p-n junction was demonstrated by Nakamura et al. [11, 12], as shown in Fig. 15.8. They used a separate confinement heterojunction (SCH) structure in conjunction with a third-order Bragg grating, in a 50 μm wide stripe geometry laser. The length of the active region was 700 μm . An unpumped waveguide, of 2–3 mm length, was continuous with the active region in order to isolate the rear cleaved surface and thereby insure that lasing occurred solely because of distributed feedback rather than mirror reflection. In the structure shown in Fig. 15.8, injected electrons are confined to the 0.2 μm thick active layer by the 0.1 μm thick layer of p-Ga_{0.83}Al_{0.17}As, while photons of the optical mode spread to the interface with the p-Ga_{0.93}Al_{0.07}As layer. Thus, the optical mode is perturbed by the grating structure, but the active region (in which recombination occurs) is far enough away from the grating so that nonradiative recombination due to lattice damage is minimized. Devices of this type were operated cw at room temperature, exhibiting threshold current densities of about $3.5 \times 10^3 \text{ A/cm}^2$. For a grating spacing of 3684 Å, a single longitudinal mode was observed with a wavelength of 8464 Å.

The concept of a separate confinement heterostructure (SCH) laser [11, 12] has been further developed and extended to InGaAsP lasers emitting at wavelengths of 1.3 μm and 1.55 μm . For example Tsuji et al. [13] have produced the SCH buried-heterostructure laser diode shown in Fig. 15.9, which emits at 1.55 μm . This device features a first-order grating of 0.234 μm spacing fabricated by holographic photolithography. A buried-heterojunction stripe geometry was used for current confinement and for transverse mode control. A transparent InP window structure served to isolate one of the end facets to suppress Fabry-Perot oscillation. It should be noted that this type of window structure also suppresses a threshold degeneracy in which two longitudinal modes closely spaced in frequency about the ω_m predicted by (15.22) both tend to oscillate [14]. The approximate relation (15.22) does not reveal this degeneracy, which can be shown to exist only by an exact solution.

Fig. 15.8 a,b Stripe geometry DFB laser with separate optical and carrier confinement [11, 12]. **a** Device structure; **b** close-up of epilayers

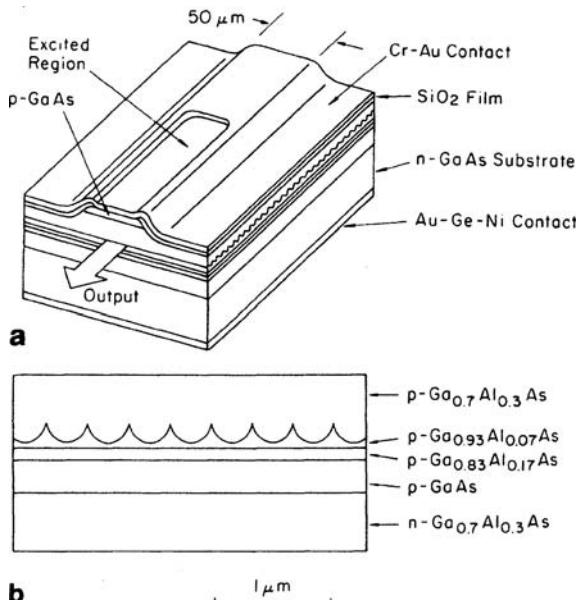
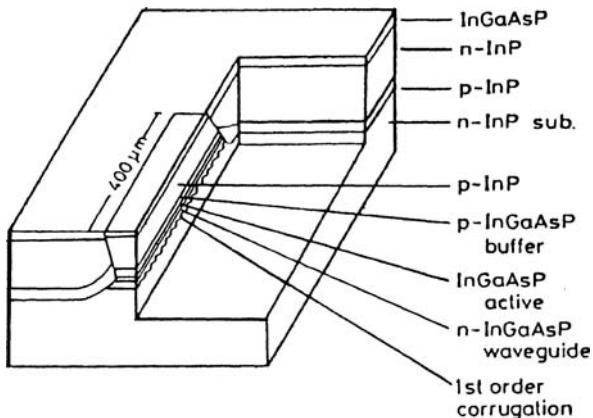


Fig. 15.9 An SCH-DFB-BH laser [11]



Lasers fabricated with the structure shown in Fig. 15.9 have been shown to have threshold current as low as 11 mA and to exhibit stable single longitudinal mode CW operation at ambient temperatures up to 106°C [13]. Lifetests on these devices have indicated only minimal degradation over thousands of hours of operation. The average output degradation rate was 0.56% per thousand hours, which is comparable with that of conventional Fabry-Perot buried-heterostructure lasers [15].

An alternative approach to separate the grating from the active light emitting region of the laser is to use the laterally-coupled geometry shown in Fig. 15.10

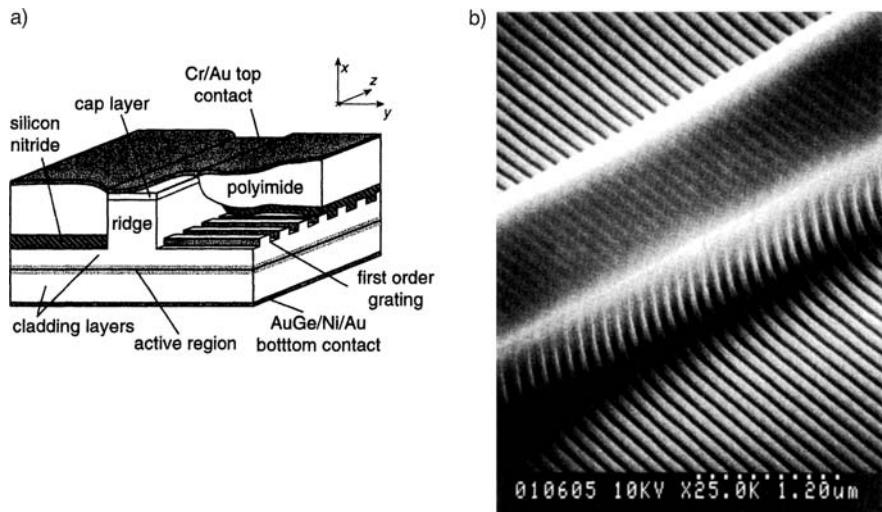


Fig. 15.10 a Diagram of a laterally-coupled distributed feedback (LC-DFB) ridge laser diode [16]. **b** Scanning electron micrograph of the first order grating pattern over the 2 μm wide ridge defined in PMMA by electron beam lithography [16]

[16]. In fabricating this device photolithography was employed to form 2 μm wide photoresist stripes on an MBE grown InGaAs-GaAs-GaAlAs graded-index-separate-confinement-heterostructure (GRINCH) wafer. Chemically assisted ion beam etching (CAIBE) was utilized to etch the stripe ridge to within 0.1 μm of the GRINCH. Then PMMA was applied to the wafer and first-order gratings were exposed by using electron-beam lithography. Following development of the resist, CAIBE was applied to transfer the gratings approximately 700 Å deep into the cap layer and upper cladding beside the ridge.

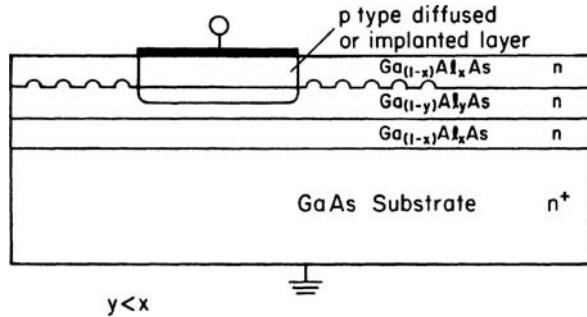
In the laterally-coupled DFB laser the two-dimensional (transverse and lateral) overlap of the evanescent field to the gratings provides the required feedback for lasing. The coupling coefficient κ is critically dependent on the ridge etch depth. Too shallow of a ridge depth reduces the transverse overlap of the field with the grating. Etching the ridge too deeply results in more index guiding of the field and reduces the lateral fill factor of the grating.

Devices of the type shown in Fig. 15.10, with cavity length of 250 μm and 1400 Å pitched gratings emitted at a wavelength of 9217 Å. Threshold current of 11 mA, differential slope efficiency of 0.46 mA/mW and single-mode output power of 15 mW were achieved.

15.2.3 DBR Lasers

The active region can also be isolated from the grating region by using the distributed Bragg reflection (DBR) structure [17] shown in Fig. 15.11. In such a device,

Fig. 15.11 Integrated DBR laser



two Bragg gratings are employed, which are located at both ends of the laser and outside of the electrically-pumped active region. In addition to avoiding nonradiative recombination due to lattice damage, placement of two grating *mirrors* outside of the active region permits them to be individually tailored to produce single-ended output from the laser. In order to achieve efficient, single-longitudinal-mode operation, one distributed reflector must have narrow bandwidth, high reflectivity at the lasing wavelength, while the other must have relatively low reflectivity for optimal output coupling.

Stoll [18] has theoretically analyzed DBR structures and shown that the transmittivity and reflectivity are given by the expressions

$$T = \frac{\gamma \exp[-i(\beta - \Delta\beta)L]}{(\alpha_p + i\Delta\beta) \sinh(\gamma L) + \gamma \cosh(\gamma L)}, \quad (15.24)$$

and

$$R = \frac{-ik \sin(\gamma L)}{(\alpha_p + i\Delta\beta) \sinh(\gamma L) + \gamma \cosh(\gamma L)}, \quad (15.25)$$

where $\gamma^2 \equiv \kappa^2 + (\alpha_p + i\Delta\beta)^2$, and where L is the length of the distributed reflector, and α_p is the distributed loss coefficient in the waveguide passive (unpumped) regions. The coupling coefficient κ and the phase constant deviation from the Bragg condition $\Delta\beta$; are as previously defined. Stoll [18] has also shown that the longitudinal mode spacing between the m th and $(m \pm 1)$ modes of a DBR laser is given to good approximation by the expression

$$\Delta \equiv |\beta_m - \beta_{m\pm 1}| = \frac{\pi}{L_{\text{eff}}}, \quad (15.26)$$

where the effective cavity length L_{eff} is given by

$$L_{\text{eff}} = L_a \left(1 + \frac{L_2}{2L_a(\alpha_p L_1 + 1)} + \frac{1}{2L_a \left(\alpha_p + \sqrt{k^2 + \alpha_p^2} \right)} \right). \quad (15.27)$$

The parameters L_1 and L_2 are the lengths of the two distributed reflectors, and L_a is the length of the active region. Single-ended optical power emission is assumed to occur from the low-reflectivity DBR of length L_2 . Typically, L_1 is on the order of one or two mm, while L_2 is several hundred micrometers. For example, for the case of $\alpha_p = 1 \text{ cm}^{-1}$, $L_a = 275 \mu\text{m}$, $k = 20 \text{ cm}^{-1}$, $L_1 = 1.47 \text{ mm}$, $L_2 = 240 \mu\text{m}$, (15.27) yields $L_{\text{eff}} = 501 \mu\text{m}$, thus leading to a mode spacing of $\Delta = 63 \text{ cm}^{-1}$.

Because they can be conveniently optimized for single-ended output at a pre-selected wavelength (determined by the grating spacing), DBR lasers are well suited to OIC frequency multiplexing applications, in which a number of sources operating at well-defined wavelengths are desired. For example, Lee et al. [19] have fabricated 21 DBR lasers on one InP chip by MOCVD of GaInAsP layers. The grating spacing of each laser was chosen to give a wavelength separation of 8 Å, with the array centered at about 1.3 μm wavelength.

DBR lasers have also been fabricated in InGaAsP/InP to produce emission at 1.55 μm wavelength. For example, Kaiser et al. [20] have integrated a tunable DBR laser, butt-coupled to a directional coupler on an InP substrate by using selective area epitaxy and strip-loaded waveguides. Average coupling efficiencies of 52% and maximum values of 63% were achieved at power levels greater than 1 mW. Tsang et al. [21] have reported InGaAs/InGaAsP multiquantum-well DBR lasers fabricated by chemical beam epitaxy with CW threshold currents of 10–15 mA, slope efficiency as high as 0.35 mW/mA and side mode suppression ratio of 58.5 dB. As many as 40 DBR lasers have been formed on one chip [22]. Schweizer et al. [23] have produced laterally coupled DFB laser diodes in InGaN/GaN. These devices were electrically pumped and lasing was observed at temperature up to 70°C. Because of the wider bandgap of the InGaN/GaN system the wavelength of emission was close to 400 nm.

15.3 Performance Characteristics

DFB and DBR lasers have unique performance characteristics that give them distinct advantages over conventional reflective-end-face lasers in many discrete-device applications, as well as in OIC's.

15.3.1 Wavelength Selectability

In cleaved-end-face lasers, the light emission occurs at a wavelength determined jointly by the gain curve and the modal characteristics of the laser. Lasing occurs for the mode (or modes) that have the highest gain. When the laser is pumped well

above the threshold level, usually a number of longitudinal modes lase simultaneously. It is very difficult, if not impossible, to obtain single mode oscillation.

In the case of DFB or DBR lasers, the emission wavelength is, of course, affected by the gain curve of the laser, but it is primarily determined by the grating spacing λ , as given by (15.3). The spacing between the l th and the $l \pm 1$ modes is generally so large compared to the linewidth of the laser gain curve that only one mode has sufficient gain to lase. Thus single-longitudinal-mode operation is obtained relatively easily in distributed feedback lasers. This gives them a distinct advantage over reflective-end-face lasers in many applications. The theoretically expected controllability of emitted wavelength was demonstrated by Nakamura et al. [24] in the first operational DFB lasers, which were optically pumped devices. The set of laser samples included both SH devices, with a $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ confining layer, and reduced-carrier concentration waveguide lasers, with a guiding layers dopant concentration of $n = 6 \times 10^{16} \text{ cm}^{-3}$ on a substrate with $n = 1 \times 10^{18} \text{ cm}^{-3}$. Both types of devices were cooled to 77 K, and were optically pumped by means of a Rhodamine B dye laser emitting at 6300 Å wavelength. Distributed feedback was provided by a third-order Bragg grating formed at the surface of the semiconductor lasers by ion milling through a photoresist mask to a depth of 500 Å. The dependence of lasing wavelength on grating spacing (in different samples) is shown in Fig. 15.12. Good agreement with the values predicted by (15.3) was obtained for a waveguide effective index of refraction $n_g = 3.59$. It was possible to controllably select laser emission wavelengths over a range of 45 Å by changing the grating period in the range from 3450 to 3476 Å. The controllability of emission wavelength makes DFB lasers particularly useful in wavelength multiplexing applications. The basic wavelength selectability of the DBR laser, which permits channel spacing on the order of 10 Å, can be further improved to < 10 Å by using a sampled grating distributed Bragg reflector (SGDBR) [19].

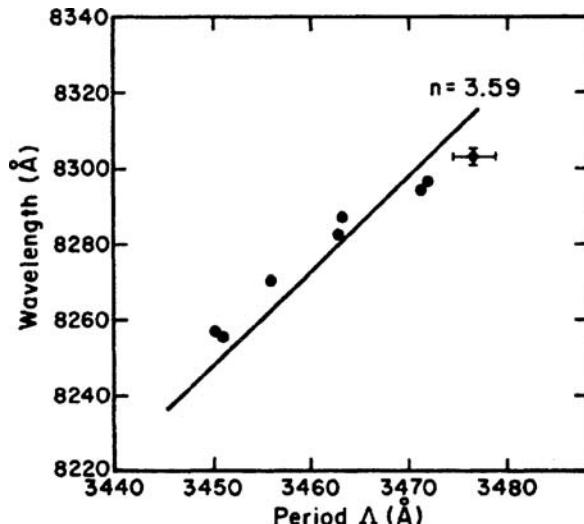


Fig. 15.12 Dependence of laser emission wavelength on grating spacing in a DFB laser [24]

As mentioned previously, there is a fundamental oscillation degeneracy in which two longitudinal modes of frequencies close to that predicted by (15.22) both have an equal probability of lasing. This degeneracy can be lifted either by providing a window structure in front of one facet as in Figs. 15.8 and 15.9, or by forming a grating with a quarter wavelength phase shift between two equal uniform corrugation sections [25, 26].

Fine tuning of the grating-selected wavelength can be accomplished by angling the active stripe at an oblique angle with respect to the axis of the grating bars [27]. A shift of approximately 170 Å was produced by angling the stripe at 7.7°, while an angle of 3.7° resulted in a shift of about 60 Å, in lasers operating in the 1.52–1.54 μm range.

15.3.2 *Optical Emission Linewidth*

In addition to providing a means of accurately selecting the peak emission wavelength, grating feedback also results in a narrower linewidth of the optical emission. The spectral width of the emission line is established by a convolution of the laser gain curve with the mode-selective characteristics of the laser cavity. Since the grating is much more wavelength selective than a cleave or polished end-face, the resulting emission linewidth of a DFB or DBR laser is significantly less than that of reflective-end-face laser. The single-mode linewidth of a conventional cleaved-end-face laser is typically about 1 or 2 Å, as described in Chapter 12 (about 50 GHz). However, values reported for modern DFB and DBR lasers with sophisticated grating structures range from about 50–100 kHz.

For example, Okai et al. [28] demonstrated a linewidth of less than 98 kHz in a three-section corrugation-pitch-modulated multiple-quantum-well DFB laser over a tuning range of 1.3 nm, at a power level of 10 mW. Mawatari et al. [29] reported a linewidth of 80 kHz at a power level of 40 mW in a modulation-doped, strained-multiple-quantum-well DFB laser, and Okai et al. [30] achieved a linewidth of only 56 kHz at a power level of 26 mW in a corrugation-pitch-modulated MQW DFB laser. Progress has continued, so that some commercially available DFB laser diodes have now been measured to have linewidths at 1.55 μm of less than 25 Å [31, 32].

The narrower linewidth obtainable with distributed feedback lasers is particularly important in optical communications applications, because the modulation bandwidth is ultimately limited by the linewidth of the laser source, and so is the density of WPM channels.

15.3.3 *Stability*

The stability of emitted wavelength and threshold current density during operation under varying ambient conditions are of key importance in many applications. Variation of emitted wavelength with junction temperature is a serious problem in

conventional cleaved-end-face lasers, with drifts of 3 or 4 Å per degree centigrade typically being observed [11, 12]. This drift is extremely detrimental to the operation of optical communication systems in which narrow bandpass optical filters are used to reduce background light intensity reaching the receiver, and thereby improve signal-to-noise ratio. A drift of as little as 10 Å might put the wavelength of the optical beam outside the bandpass of the filter. Not only is the long term drift of wavelength with ambient temperature important, but also transient drift or *chirp* caused by junction heating during pulsed operation or modulation must be considered.

Distributed feedback lasers offer improved wavelength stability as compared to cleaved-end-face lasers, because the grating tends to lock the laser to a given wavelength. For example, Nakamura et al. [11, 12] have measured the temperature dependence of emitted wavelength for both cleaved-end-face and DFB lasers fabricated from the same wafer. The data are shown in Fig. 15.13. The length of the cleaved laser was 570 μm, while that of the DFB laser was 730 μm. The grating spacing in the DFB laser was 3814 Å. Both devices were fabricated from the same DH GaAlAs wafer. Note that the cleaved laser exhibited a wavelength drift of 3.7 Å/°C, while the

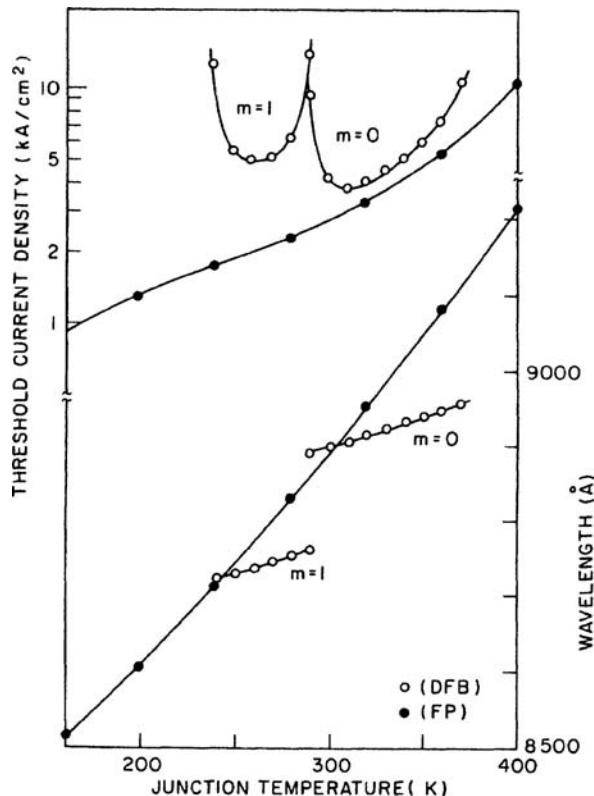


Fig. 15.13 Threshold current density and lasing wavelength as functions of junction temperature [11, 12]

DFB laser drifted only $0.8 \text{ \AA/}^{\circ}\text{C}$, over temperature spans of about 100° . However, the DFB laser jumped from the $m = 0$ mode to the $m = 1$ mode at about 280 K. From the threshold current density data shown in Fig. 15.13, it can be seen that the DFB laser required roughly the same current as the cleaved laser when oscillating in the $m = 0$ mode. However, its J_{th} was about a factor of three larger for the $m = 1$ mode. Also the J_{th} of the DFB laser increased sharply at the temperatures where a transition between modes occurred. These rapid increases resulted from mismatch between the laser gain curve and the modal resonances (established by the grating period) at those temperatures. Martin et al. [16] have made similar measurements of wavelength drift for the laterally-coupled DFB laser of Fig. 15.10, in the temperature range from 5 to 50°C . A wavelength shift of $0.63 \text{ \AA/}^{\circ}\text{C}$ was observed for the LC-DFB laser, while a shift of $2.8 \text{ \AA/}^{\circ}\text{C}$ was measured for a standard Fabry-perot cleaved-endface ridge laser of the same heterostructure and dimensions as the DFB device. Even high-power DFB lasers exhibit low thermal drift when they are properly designed. For example, Takigawa [31] has reported as 50 mW 780 nm GaAlAs DFB laser which operates in a stable single longitudinal mode with a lasing wavelength temperature coefficient of $0.65 \text{ \AA/}^{\circ}\text{C}$. In that case a buried twin-ridge heterostructure was used to produce a stable fundamental spatial mode.

The improved temperature stability of distributed feedback lasers results from the fact that the shift of emitted wavelength in the cleaved laser follows the temperature dependence of the energy bandgap, while the shift in wavelength of the DFB laser follows only the temperature dependence of the index of refraction. This improved stability of the DFB laser makes it very useful in applications where wavelength filtering or wavelength division multiplexing are employed. Hence, even in many cases where a discrete laser is being used rather than an OIC, the DFB or DBR laser is to be preferred over a Fabry-Perot end-face laser. DFB and DBR lasers have now become the standards for long-distance transmission in the telecommunications industry.

15.3.4 Commercially Available DFB Lasers

The results of research and development of DFB lasers have been transferred to the commercial marketplace, where there are many suppliers. DFB lasers are now commercially available with emission wavelengths throughout the near-infrared range of 730–2800 nm. Output power levels range from about 10 to 200 mW. Linewidths are typically 2–4 mHz, but at certain wavelengths, linewidth as narrow as 100 kHz is available.

15.4 Nanoscale DFB Lasers

As the techniques of semiconductor microfabrication have improved to the point of becoming nanofabrication, it has become possible to make some new Bragg-type reflectors.

15.4.1 Semiconductor Air Bragg Reflector Lasers

A number of different research groups have demonstrated that it is possible to make Bragg-type reflectors by machining slots into the semiconductor surface [33–38]. These narrow air-gaps between “walls” of semiconductor material can form Bragg reflectors if their dimensions are small enough. Because of the large difference in index of refraction between the semiconductor and air the reflections at each interface are relatively strong compared to those at each bar of a conventional Bragg grating. Thus only a few bars are required to get essentially 100% reflection. In a conventional grating formed in the semiconductor, thousands of bars are required. Thus these high-reflectivity semiconductor/air Bragg reflectors (SABAR) can be used to make short-cavity, high-reflectance lasers.

Mukaihara et al. [38] have made a laser of this type in InGaAsP by MOCVD and integrated it with a waveguide, as shown in Fig. 15.14.

Within the Bragg reflector the thickness of the semiconductor layers is $7/4\lambda$ and the thickness of the air gaps is $3/4\lambda$. The structure was fabricated by using e-beam lithography to mask the surface and wet etching to produce the vertical groove. The photomicrographs in Fig. 15.14 show the shape of the reflecting bars and also the zone in which the laser active region is coupled to the waveguide. This laser emits at a wavelength of $1.3 \mu\text{m}$. The threshold current of this ridge waveguide laser with SABAR (cavity length $L = 160 \mu\text{m}$, ridge width $W = 7 \mu\text{m}$, SABAR pairs $N = 3$) is 20 mA. The misalignment of the laser active region and the waveguide in this

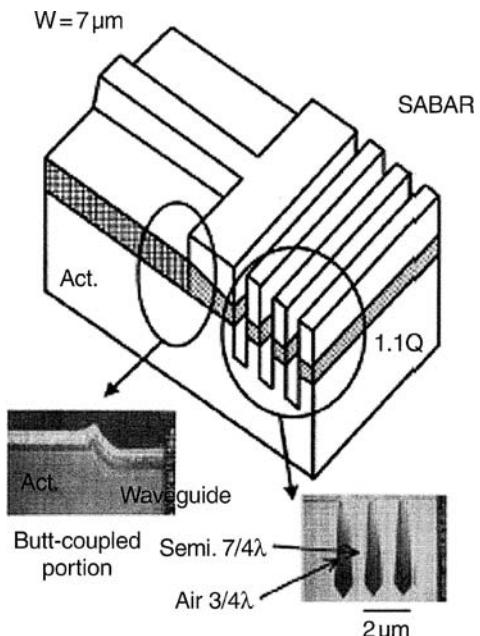


Fig. 15.14 Integrated waveguide and laser with SABAR reflector [38] © 1999 IEEE

sample introduced coupling loss and increased the threshold current. The authors calculate that a threshold current of $2 \text{ mA}/\mu\text{m}$ (of width) should be attainable.

Hofling et al. [37] have fabricated ridge-waveguide, edge-emitting, GaInAs-AlGaAs microlasers with SABAR type reflectors, having $N = 2$ air/semiconductor pairs at each endface forming third-order gratings. The gratings were formed by reactive ion etching (RIE). Reflectivities greater than 75% were achieved. The lasers had threshold currents as low as 2 mA for an 80 μm cavity length and 6 mA for a 40 μm length. Cavity lengths as short as 28 μm still produced lasing. The small size of lasers with SABAR mirrors is a key advantage when they are considered for integration into DWDM transmitter chips. Note that a third-order grating was used in this case, which is easier to fabricate than a first-order grating, but has lower efficiency. Second-order gratings are not used because of their large surface-emission component, as described in Section 15.1.2.

15.4.2 Quantum Dot DFB Lasers

A quantum dot (QD) is a semiconductor structure with all three spatial dimensions less than 10 nm. Electrons, holes, and excitons are confined by the dot structure, which can significantly change the behavior of the material in which the quantum dots are formed. Quantum dots are discussed in Chapter 22. For additional information on quantum dots see Michler [39] or Schweizer et al. [40].

By introducing quantum dots into the grating region of a DFB laser it is possible to improve the operating characteristics of the device. For example, Kim et al. [41] have grown self-assembled InAs/InAlGaAs QDs on InP/InGaAs grating structures by molecular beam epitaxy. Ridge-waveguide QD DFB lasers with a stripe width of 3 μm were fabricated. Single-mode cw lasing at a wavelength of 1.56 μm was achieved at room temperature, and pulsed-mode lasing was observed up to 70°C.

Su and Lester [42] have examined the dynamic properties of distributed feedback lasers based on InAs/InGaAs quantum dots. They found that the linewidth of the QD DFBs was more than one order of magnitude narrower than that of conventional quantum well (QW) DFBs at comparable output powers.

Kamp et al. [43] have made DFB QD lasers based on a single layer of InGaAs-AlGaAs self-organised quantum dots grown by molecular beam epitaxy, with metal gratings patterned laterally in a ridge waveguide to provide feedback. They observed threshold current = 14 mA, differential efficiency = 0.33 W/A, and side-mode suppression ratios greater than 50 dB.

Problems

- 15.1 What is the grating spacing required for a DFB laser to operate at a wavelength $\lambda_0 = 8950 \text{ \AA}$ in GaAs if a first-order grating is desired?
- 15.2 Repeat Problem 15.1 for the case of a laser formed in an active region of $\text{Ga}_{0.8}\text{Al}_{0.2}\text{As}$.

- 15.3 If the grating of Problem 15.1 is fabricated by using the technique of Fig. 7.10, determine two combinations of prism material and angle (α) that could be used with a He-Cd laser ($\lambda_0 = 3250 \text{ \AA}$).
- 15.4 (a) If a surface grating is formed on a waveguide, what is the grating spacing λ required for fourth-order Bragg diffraction, in terms of λ and n_g ?
- (b) Sketch the reflected wave directions, and indicate the angles between them.
- (c) In general, can a scattered wave perpendicular to the direction of the waveguide be observed for odd orders, such as the first and third order, of Bragg diffraction?
- 15.5 A buried heterostructure, multiple contact, distributed feedback (DFB) laser has a grating periodicity of 1200 \AA and a waveguide index of refraction equal to 3.8. The interaction length of the grating is $500 \mu\text{m}$. Assume that only a single mode oscillates and that feedback comes from a first-order Bragg reflection. In this problem you may neglect the slight shift in wavelength which occurs in a DFB laser and assume that emission is at the Bragg wavelength.
- (a) What is the emitted wavelength of the laser?
- (b) If current is injected into the modulator section of the diode so as to change the waveguide index by one part in 10^7 , what will be the resulting change in laser emitted wavelength (frequency)?
- 15.6 A high-gain, distributed feedback (DFB), end-face emitting laser is fabricated in GaInAsP such that its effective guide index of refraction is 3.65. It has a first-order grating of periodicity $= 1150 \text{ \AA}$ and length $300 \mu\text{m}$.
- (a) What is the frequency of the lowest-order mode in rad/s?
- (b) What is the mode spacing in rad/sec between the lowest-order mode and the next higher-order mode?
- 15.7 (a) Why are second order gratings not used in conventional edge emitting DFB lasers?
- (b) Describe a type of laser in which a second order grating would be especially useful.
- 15.8 What is the mode frequency spacing (in mHz) in a DFB laser with a guide index of 3.62 and a length of $250 \mu\text{m}$?

References

1. Z.G. Pinsker: *Dynamical Scattering of X Rays in Crystals*, Springer Ser. Solid-State Sci., Vol. 3 (Springer, Berlin, Heidelberg 1978)
2. B.K. Agarwal: *X-Ray Spectroscopy*, 2nd edn., Springer Ser. Opt. Sci., Vol. 15 (Springer, Berlin, Heidelberg 1991)
3. H. Kogelnik, C.V. Shank: *J. Appl. Phys.* **43**, 2327 (1972)

4. A. Yariv: IEEE J. **QE-9**, 919 (1973)
5. S. Wang: IEEE J. **QE-10**, 413 (1974)
6. D.R. Scifres, R.D. Burnham, W. Streifer: Appl. Phys. Lett. **26**, 48 (1975)
7. W. Streifer, D.R. Scifres, R.D. Burnham: IEEE J. **QE-11**, 867 (1975)
8. A. Yariv: *Optical Electronics*, 4th edn. (Holt, Rinhart, Winston, New York 1991) p. 503
9. R. Buchmann, H. Dietrich, G. Sasso, P. Vettiger: Microelectron. Eng. **9**, 485 (1989)
10. D.R. Scifres, R. Burnham, W. Streifer: Appl. Phys. Lett. **25**, 203 (1974)
11. M. Nakamura, K. Aiki, J. Umeda, A. Yariv: Appl. Phys. Lett. **27**, 403 (1975)
12. K. Aiki, M. Nakamura, J. Umeda: IEEE J. **QE-12**, 597 (1976)
13. S. Tsuji, A. Ohishi, N. Nakamura, M. Hirao, N. Chinone, H. Matsumura: IEEE J. LT-**5**, 822 (1987)
14. P. Bhattacharya: *Semiconductor Optoelectronic Devices* (Prentice-Hall, Englewood Cliffs, NJ 1994) pp. 286–292
15. S. Tsuji, K. Mizushima, M. Hirao, M. Nakamura: IEEE Int'l Conf. on Communications, Amsterdam, The Netherlands (1984) Proc. p. 1123
16. R. Martin, S. Forouhar, S. Keo, R. Lang, R.G. Hunsperger, R. Tiberio, P. Chapman: Electron. Lett. **30**, 1058 (1994)
17. S. Wang: IEEE J. **QE-10**, 413 (1974)
18. H.M. Stoll: IEEE Trans. CAS-**26**, 1065 (1979)
19. S.L. Lee, I.F. Jang, C.Y. Wang, C.T. Pien, T.T. Shih: Monolithically integrated multi-wavelength sampled grating DBR lasers for dense WDM applications. IEE J. Selected Topics Quant. Electron. **6**, 197 (2000)
20. R. Kaiser, F. Fidorra, H. Heidrich, P. Albrecht, W. Rehbein, S. Malchow, H. Schroeter-Janssen, D. Franke, G. Sztefka: 6th Int'l Conf. on InP and Related Materials. Santa Barbara, CA (1994) Proc. p. 474
21. W. Tsang, M. Wu, Y. Chen, F. Choa, R. Logan, S. Chu, A. Sergent, P. Magill, K. Reichmann, C. Burrus: IEEE J. **QE-30**, 1370 (1994)
22. K. Kudo, M. Ishizaka, T. Sasaki, H. Yamazaki, M. Yamaguchi: 1.52–1.59 μm range different-wavelength modulator-integrated DFB-LD's fabricated on a single wafer. IEEE Photon Tech. Lett. **10**, 929 (1998)
23. H. Schweizer, H. Gräbeldinger, V. Dumitru, M. Jetter, S. Bader, G. Brüderl, A. Weimar, A. Lell, V. Härlé: Laterally coupled InGaN/GaN DFB laser diodes, Physica Status Solidi (a) **192**, 301 (2002)
24. M. Nakamura, H.W. Yen, A. Yariv, E. Garmine, S. Somekh, H.L. Garvin: Appl. Phys. Lett. **23**, 224 (1973)
25. H.A. Haus, C.V. Shank: IEEE J. **QE-12**, 532 (1976)
26. S. Akiba, Y. Matsushima, M. Usami, K. Utaka: Electron. Lett. **23**, 316 (1987)
27. W. Tsang, R. Kapre, R. Logan, T. Tanbun-Ek: IEEE Photon. Tech. Lett. **5**, 978 (1993)
28. M. Okai, T. Tsuchiya: Electron. Lett. **29**, 349 (1993)
29. H. Mawatari, F. Kano, N. Yamamoto, Y. Kondo, Y. Tohmori, Y. Yoshikuni: Jpn. J. Appl. Phys. Pt. 1. **33**, 811 (1994)
30. M. Okai, A. Tsuchiya, A. Takai, N. Chinone: IEEE Photon. Tech. Lett. **4**, 526 (1992)
31. S. Takigawa, T. Uno, M. Kume, K. Hamada, N. Yoshikawa, H. Shimizu, G. Kano: IEEE J. **QE-25**, 1489 (1989)
32. G. Giuliani, M. Norgia: Diode laser linewidth measurement by means of self-mixing interferometry. IEEE Phot. Tech. Lett. **12**, 1028 (2000)
33. K.C. Shin, M. Tamura, A. Kasukawa, N. Serizawa, S. Kurihashi, S. Tamura, S. Arai: Low threshold current density operation of GaInAsP-InP laser with multiple reflector microcavities. IEEE Photon. Tech. Lett. **7**, 1119 (1991)
34. T. Baba, M. Hamasaki, N. Watanabe, P. Kaewplung, A. Matsutani, T. Mukaihara, F. Koyama, K. Iga: A novel short-cavity laser with deep-grating distributed Bragg reflectors, Jpn. J. Appl. Phys. **35**, 1390 (1996)

35. R. Jambunathan, J. Singh: Design studies for distributed Bragg reflectors for short-cavity edge-emitting lasers. *IEEE J. Quant. Electron.* **33**, 1180 (1997)
36. S. Oku, T. Ishii, R. Iga, T. Hiroto: Fabrication and performance of AlGaAs-GaAs distributed Bragg reflector lasers and distributed feedback lasers utilizing first-order diffraction gratings formed by a periodic groove structure. *IEEE J. Selected Topics Quant. Electron.* **5**, 682 (1999)
37. E. Hofling, F. Schafer, J. Reithmaier, A. Forchel: Edge-emitting GaInAs-AlGaAs microlasers, *IEEE Photon. Tech. Lett.* **11**, 943 (1999)
38. T. Mukaihara, N. Yamanaka, N. Iwai, M. Funabashi, S. Arakawa, T. Ishikawa, A. Kasukawa: Integrated GaInAsP laser diodes with monitoring photodiodes through semiconductor/air Bragg reflector (SABAR). *IEEE J. Selected Topics Quant. Electron.* **5** 469 (1999)
39. P. Michler, (ed.): *Single Quantum Dots-Fundamentals, Applications and New Concepts*, Springer Topics in Applied Physics Series, vol. 90 (Springer, Berlin, Heidelberg, 2003)
40. H. Schweizer, M. Jetter, F. Scholz: *Quantum-Dot Lasers*, Springer Topics in Applied Physics Series, vol. 90 (Springer, Berlin, Heidelberg, 2003)
41. J.S. Kim, C-R. Lee, H-S. Kwack, B.S. Choi, E. Sim, C.W. Lee, D.K. Oh: 1.55 μm InAs/InAlGaAs quantum dot DFB lasers, *IEEE Trans. Nanotechnol.* **7**, 128 (2008)
42. H. Su, L.F. Lester: Dynamic properties of quantum dot distributed feedback lasers: high speed, linewidth and chirp, *J. Phys. D: Appl. Phys.* **38**, 2112 (2005)
43. M. Kamp, M. Schmitt, J. Hofmann, F. Schafer, J.P. Reithmaier, A. Forchel: InGaAs/AlGaAs quantum dot DFB lasers operating up to 213°C, *Electron. Lett.* **35**, 2036 (1999)

Chapter 16

Direct Modulation of Semiconductor Lasers

In Chapters 9 and 10 techniques were described for modulating the light of a semiconductor laser by using external electro-optic or acousto-optic modulators. However, it is also possible to internally modulate the output of a semiconductor laser by controlling either the current flow through the device or some internal cavity parameter. Such direct modulation of the laser output has the advantages of simplicity and potential for high frequency operation. The topic of direct modulation of injection lasers is considered in this chapter; this follows the discussions of semiconductor laser and amplifier fundamental principles and operating characteristics in Chapters 11, 12, 13 and 14, so that the reader will be better prepared to appreciate the subtleties of the methods involved.

16.1 Basic Principles of Direct Modulation

The light output of a semiconductor laser can be directly modulated, i.e., made to vary in response to changes within the laser cavity, so as to produce amplitude modulation (AM), optical frequency modulation (FM), or pulse modulation (PM). Most often, the laser output is either amplitude or pulse modulated by controlling the current flow through the device. However, other parameters, such as the dielectric constant or the absorptivity of the laser cavity material can be varied to produce FM, as well as AM and PM of the output. The basic principles of these modulation techniques, most of which were developed during the late 1960's and 1970's, are discussed in this section. More sophisticated improvements, which have led to reliable direct modulation of laser diodes at microwave frequencies in later years, are covered in Section 16.2.

16.1.1 Amplitude Modulation

The basic arrangement for amplitude modulation of a laser diode by control of current flow is illustrated in Fig. 16.1. The laser diode must be dc biased to a point above the lasing threshold level in order to avoid the sharp break in the output curve

at threshold. Above threshold, the dependence of output power on current is very linear for a well made diode, as discussed in Chapter 13. The ac modulating signal must be isolated from the dc bias supply, and also the dc bias must be prevented from reaching the modulating signal source. At low frequencies, this can be accomplished by means of a simple inductor and capacitor, as shown in Fig. 16.1. Above about 50 MHz, more sophisticated high-pass and low-pass filter circuits must be used.

For the direct current modulation method shown in Fig. 16.1, the modulation depth is given by

$$\eta = \frac{P_p - P_m}{P_p}, \quad (16.1)$$

where P_p is the peak optical power and P_m is the minimum. The maximum modulation depth over which linear response can be obtained is given by

$$\eta_{\max} = \frac{P_p - P_t}{P_p}, \quad (16.2)$$

where P_t is the optical output at threshold. Because the power at threshold P_t is usually only about five or ten percent of P_p , the maximum modulation depth can, in principle, be greater than 90%. The expression (16.2) is based on the tacit assumption that the dc bias is chosen so that the operating point for zero input signal is at the center of the linear portion of the output curve, i.e., at a power of $(P_p - P_t)/2$.

The amplitude modulation process can be described in terms of a pair of nonlinear rate equations. As given by Lasher [1], these relations state that

$$\frac{dN_e}{dt} = \frac{1}{eV} - \frac{N_e}{\tau_{sp}} - GN_{ph}, \quad (16.3)$$

and

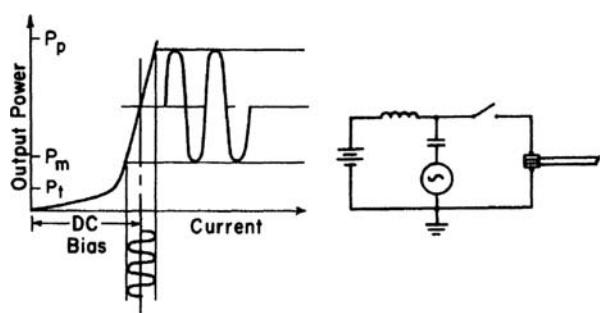


Fig. 16.1 Basic bias circuit and output characteristics for directly AM modulated laser diode

$$\frac{dN_{ph}}{dt} = \left(G - \frac{1}{\tau_{ph}} \right) N_{ph}, \quad (16.4)$$

where N_e is the electron inversion number, N_{ph} is the number of photons, I is the current, V is the volume of the active region, τ_{sp} is the spontaneous electron lifetime. τ_{ph} is the photon lifetime, and G is the stimulated emission rate. The relations (16.3) and (16.4) assume a single lasing mode, into which there is negligible spontaneous emission. For small-signal analysis, a small time varying signal $I(t)$ is imposed on the dc bias current I_{dc} . In the conventional small signal approximation, the resulting small variations in the photon number $n_{ph}(t)$ and the electron inversion number $n_e(t)$ about their average values \bar{N}_{ph} and \bar{N}_e are given by [2]

$$\frac{d^2}{dt^2} \left\{ \frac{n_e}{n_{ph}} \right\} + \gamma \frac{d}{dt} \left\{ \frac{n_e}{n_{ph}} \right\} + \omega_0^2 \left\{ \frac{n_e}{n_{ph}} \right\} = \left\{ \frac{\frac{1}{eV} \frac{dI(t)}{dt}}{\frac{g\bar{N}_{ph}I(t)}{eV}} \right\}, \quad (16.5)$$

where

$$\omega_0^2 = \frac{1}{\tau_{sp}} \left(gN_0 + \frac{1}{\tau_{ph}} \right) \left(\frac{I_{dc}}{I_{th}} - 1 \right), \quad (16.6)$$

and

$$\gamma = \frac{1}{\tau_{sp}} + \tau_{ph}\omega_0^2. \quad (16.7)$$

In (16.5) and (16.6), the stimulated emission rate is assumed to be given by

$$G = g(N_e - N_0), \quad (16.8)$$

where N_0 is the inversion required to overcome bulk losses, and g is a constant of proportionality. If sinusoidal modulation of the current is used such that

$$I(t) = I_m \cos \omega_m t, \quad (16.9)$$

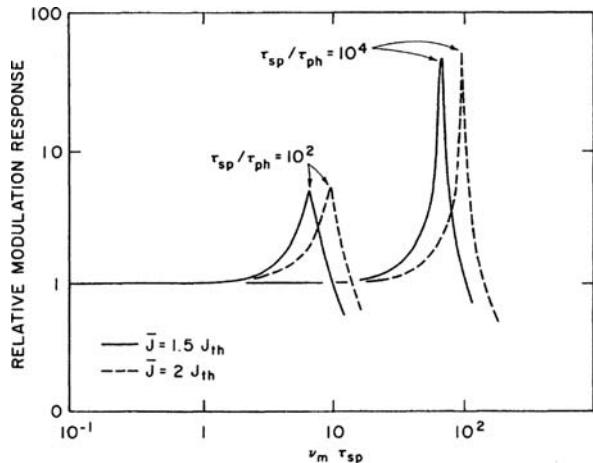
then the modulation depth is given by [2]

$$\eta = \frac{n_{ph}(\omega_m)}{\bar{N}_{ph}} = \frac{\frac{gI_m}{eV}}{\omega_0^2 - \omega_m^2 + i\omega_m\gamma}. \quad (16.10)$$

The expression (16.10) exhibits a pronounced peak at a frequency ν_{max} , given by

$$\nu_{max} = \frac{\omega_{max}}{2\pi} = \frac{1}{2\pi} \left(\omega_0^2 - \frac{\gamma^2}{2} \right)^{1/2}, \quad (16.11)$$

Fig. 16.2 Theoretical dependence of modulation depth on modulation frequency [3]. The curve has been calculated based on typical parameters for a GaAs laser and has been normalized to the value of η for ω_m near zero. Curves are plotted for two values of τ_{sp}/τ_{ph} to illustrate the strong effect of changing the Q of the laser cavity



as shown in Fig. 16.2, in the theoretical results produced by Ikegami and Suematsu [3], who have also observed this peak experimentally. For conventional GaAs laser diodes v_{max} is on the order of several gigahertz. The existence of such a peak, followed by a rapid decrease in response of course, implies an upper limit to modulation frequency. However, significant effects that are not included in the simple model presented thus far permit specially designed laser diodes to operate at much higher frequencies. These effects are discussed in detail in Section 16.2.

16.1.2 Pulse Modulation

Pulse modulation is particularly convenient in semiconductor lasers because of their very short turn-on and turn-off times when the pumping current is pulsed. Pulse rise and fall times on the order of a few hundred picoseconds, typical of DH stripe geometry lasers, permit subnanosecond pulses to be generated, with nanosecond spacing [4]. However, the laser must be biased to just below threshold in order to obtain such high speed pulsing. Otherwise, there is an initial delay between the application of the current pulse and the emission of the light pulse, which is given by [5]

$$t_d = \tau_{sp} \ln [I_p / (I_p - I_{th})], \quad (16.12)$$

where I_p is the peak pulse current. This delay vanishes when the laser is biased to I_{th} .

The maximum permissible duty cycle of a particular laser can also impose a limit to high pulse-repetition-rate operation. In the case of DH stripe geometry lasers, capable of cw operation at room temperature, duty cycle is not a problem. However,

in less expensive SH or homojunction lasers, the maximum duty cycle must not be exceeded, or else junction heating will result in wavelength drift and increased threshold current. When duty cycle is a limiting factor, the maximum data rate can be made larger than the pulse repetition rate by using a coding that permits each pulse to carry more than one bit. For example, a 1 ns pulse at an average repetition rate of 30 MHz can transmit a data rate of 150 Mbit/s if pulse-interval modulation is used [6].

Laser diodes can also be pulse-width modulated by controlling the width of the drive current pulse. Alternatively, Fenner [7] has proposed that the initial delay between the application of the current pulse and the onset of light emission can be used to produce pulse-width modulation. From (16.12) it can be seen that I_d is a strong function of I_p . Hence, if the driving current pulses are amplitude modulated, the resulting light pulses will be pulse-width modulated. Since the relationship between I_p and t_d is nonlinear, suitable compensation must be provided in the decoding network if this method of pulse-width modulation is used.

Trains of very narrow pulses, at high pulse repetition rates, can be generated in a semiconductor diode laser by taking advantage of the phenomenon of self-pulsing. Note that the form of (16.5), for the case of $I(t) = 0$, implies that there exists the possibility of relaxation oscillations at a frequency ω_R , given by

$$\omega_R = \omega_0^2 - \left(\frac{\gamma}{2}\right)^2, \quad (16.13)$$

when the laser is biased above threshold. This *spiking resonance* is usually considered to be detrimental to laser modulation, because it results in the peaking distortion of the AM modulation frequency response curve, as described in Section 16.1.1. However, self pulsing due to augmented relaxation oscillations can be used beneficially, in some cases, to produce narrow pulses at GHz rates. For example, D'Asaro et al. [8] have reported pulse rates between 0.5 and 3.0 GHz. The exact nature of the relationship between the damped relaxation oscillations predicted by (16.1.5) and the self sustained oscillations observed in many laser diodes is not yet thoroughly understood. A number of different models have been proposed, which are discussed in more detail with regard to high frequency modulation in Section 16.2.

Self-pulsing junction lasers can be conveniently modulated by analog pulse-position modulation. In this method, the optical pulse rate is frequency-locked to an external signal modulation of the injection current at a frequency near the resonance frequency v_r , or one of its harmonics [9]. If the frequency of the locking signal is made to vary as the derivative of the information signal, the deviation of the pulse position from its average value is directly proportional to the information signal itself. Deviations of the pulse rate by as much as 10% of v_r can be obtained with only a few milliwatts of modulating microwave power [9]. Alternatively, pulse-position-modulation of self-pulsing lasers can be produced by using the current dependence of the self-induced pulse rate v_r [8], or by utilizing an external phase-locked loop to provide regenerative feedback of the self-induced oscillations of injection current that occur simultaneously with the optical oscillations [10].

All of the methods of modulating laser diodes discussed thus far rely in some way on the dependence of emitted optical power on injected current. However, it is also possible to directly modulate laser diodes by varying certain parameters of the cavity material.

16.1.3 Frequency Modulation

Direct modulation of the optical frequency of a diode laser can be produced by varying the dielectric constant ϵ of the cavity. The two-section lasers described as pulse modulated devices in Section 16.1.2 can also be operated in an FM mode [11]. In that case, the current passing through the absorber section of the dual diode is varied in order to produce a change in average cavity ϵ . Of course, the dual diode must be operated outside of the range of bias conditions for which self-pulsing oscillation can occur, even so, some **AM** modulation will be present on the output because of variations in average gain produced by the changing current in the absorber section.

An improved version of the two-section laser is the cleaved-coupled-cavity (C^3) laser, in which not only the electrical contacts but also the laser cavity itself are divided in two [12]. The C^3 laser is fabricated by first making a standard stripe-geometry Fabry-Perot laser with a cavity length of approximately 250 μm . This device is then recleaved near the middle to produce two coupled laser cavities. These cavities are only separated by a few micrometers and are held in alignment for mounting by a relatively thick electroplated contact on one side. In operation, one section of the laser is pumped above threshold by an injected current I_1 , while the dielectric constant of the second section is varied by a modulating current I_2 [13]. By controlling I_2 it is possible both to tune the laser to a desired wavelength and to frequency modulate it. For example, a maximum wavelength excursion of 150 \AA with a tuning rate of 10 $\text{\AA}/\text{mA}$ has been reported for a GaInAsP C^3 laser emitting at 1.3 μm [12].

In addition to permitting frequency modulation of the laser the dual-cavity structure also results in very stable single mode operation. The individual mode spacings of the cleaved cavities are given by

$$\Delta\lambda_1 = \frac{\lambda_0^2}{2n_{\text{eff}1}L_1} \quad (16.14)$$

and

$$\Delta\lambda_2 = \frac{\lambda_0^2}{2n_{\text{eff}2}L_2}, \quad (16.15)$$

where λ_0 is the peak emission wavelength, $n_{\text{eff}1}$ and $n_{\text{eff}2}$ are the effective indices of refraction and L_1 and L_2 are the lengths of the two cavities. Because the two cavities are strongly coupled, only those modes which coincide spectrally will constructively

combine to form the resultant modes of the coupled cavity. Thus the mode spacing of the coupled cavity is given by [12]

$$\Delta\lambda_c = \frac{\Delta\lambda_1\Delta\lambda_2}{|\Delta\lambda_1 - \Delta\lambda_2|} = \frac{\lambda_0^2}{2[|n_{\text{eff}1}L_1 - n_{\text{eff}2}L_2|]}. \quad (16.16)$$

Since $\Delta\lambda_c >> \Delta\lambda_1$ or $\Delta\lambda_2$, only one mode near the peak of the gain curve survives. The laser thus is very stable when tuned to a desired frequency by the DC component of I_2 and frequency modulation can be produced by applying a small-signal AC modulation to I_2 . In addition, a stable single mode also is maintained if direct amplitude modulation is produced by varying l_1 in the usual way. For example, it has been shown that a stable, single-frequency mode can be maintained under 2 Gb/s direct modulation with error rates of less than 10^{-10} [14].

The stable tuning and modulation characteristics of GaInAsP C³ lasers make them very attractive for use in wavelength division multiplexed transmitters for long-distance telecommunications. The C³ structure can also be used to make highly stable single mode lasers in GaAlAs [15].

Frequency modulation, with negligible AM, can be achieved by using acoustic waves to produce a change in cavity ϵ [16]. Longitudinal sound waves passing through the laser diode in the direction perpendicular to the junction cause a shift of the lasing mode to a frequency v given by

$$v = v_0 + A \exp\left(-\frac{\omega_a^2 W^2}{8v_a^2}\right) \cos(\omega_a t), \quad (16.17)$$

where v_0 is the frequency of the mode in the absence of the acoustic waves, A is a constant proportional to the peak acoustic intensity, ω_a is the acoustic frequency. v_a is the acoustic velocity in the material, and W is the width of the optical mode at $1/3$ of its peak amplitude. As in the case of Bragg type acousto-optic modulators, the frequency response of the transducer used to launch the acoustic waves is presently the predominant factor limiting modulation bandwidth. Theoretical calculations predict that the maximum bandwidth of an acoustically modulated GaAs laser, with 400 μm cavity length, can be as high as 43 GHz, being limited ultimately by mode jumping [16].

16.2 Microwave Frequency Modulation of Laser Diodes

In the previous section, a glimpse was provided of the microwave frequency modulation capabilities of the injection laser. Because such modulation plays a key role in determining the upper limit to the information carrying capacity of systems employing optical carrier frequencies, it is appropriate to now review in detail the accomplishments and limitations of the various techniques that have been used (or proposed) to directly modulate diode lasers at frequencies up to 40 GHz.

16.2.1 Summary of Early Experimental Results

It was early realized that the inherently short electron and photon lifetimes of a semiconductor laser make it well suitable for applications requiring microwave frequency modulation [17, 18]. Well before the advent of laser diodes that could operate at room temperature, modulation at X-band (8–12 GHz) frequencies [19], and even as high as 46 GHz [20] was demonstrated in cryogenically cooled devices. In the five years immediately following these initial demonstrations, work continued in order to measure lifetimes and to study resonance phenomena [21–26]. In all cases, the laser diodes were cryogenically cooled (typically to 77 K), and were mounted in either microwave waveguide [19, 26] or in a coaxial line [22, 23, 25]. Modulation frequencies in the mm wave range, as high as 46 GHz, were reached, but generally with modulation depths of only a few percent [26]. When room-temperature cw lasers became available in 1970 [27, 28], further attempts at microwave modulation were directed toward the use of these devices, because cooling to 77 K presented a serious obstacle to any practical application.

During the 1970s work done on the microwave modulation of CW, room-temperature lasers demonstrated that these devices could be directly modulated at frequencies as high as 10 GHz with modulation depth of a few percent [29] and in the range from 1–3 GHz with modulation depth as great as 30% [30–36]. However, response at frequencies greater than 3 GHz was limited both by the response characteristics of the laser itself, and by the parasitic inductance, capacitance and resistance of the diode package and surrounding microwave cavity and/or transmission line.

16.2.2 Factors Limiting Modulation Frequency

The factors that can set an upper limit to the frequency at which a laser diode can be directly modulated are summarized in Table 16.1. The first four items on the list are inherent in the laser diode chip itself, while the remaining four are associated with the diode package and microwave circuitry. Of course, not all of these factors will be important in every case. However, any one of them acting alone can be sufficient to limit the maximum modulation frequency.

The optical turn-on delay associated with the buildup of the inverted carrier population was discussed in Section 16.1.2. It can be essentially eliminated by biasing the laser diode just below threshold. The optical turn-off delay is unlikely to be a

Table 16.1 Factors limiting the modulation frequency of a laser diode

Delayed turn-on/turn-off	Parasitic capacitance and inductance
Transient multimode excitation	RF leakage radiation
Relaxation oscillations	RF absorption loss
Self-sustaining oscillations	Impedance mismatch

limiting factor at frequencies up to 100 GHz since in GaAs and GaAlAs the carrier lifetime in the inverted population region is approximately 10^{-11} s [24], and the photon lifetime is even shorter [22]. However, the carrier lifetime is about a hundred times longer in noninverted material. Hence, biasing the laser close to threshold also has the effect of eliminating the relatively slowly decaying *tail* of photons that are spontaneously emitted while the laser current is below the threshold level.

Transient multimode excitation produces a broadening of the output spectrum of a pulsed laser at the onset of the current pulse [37]. The effect is particularly important at high bit rates because such transient excitation of multiple modes can affect coupling into an optical fiber, producing pulse shape distortion. In such applications a superior quality, low threshold single mode laser is recommended.

The possibility of damped relaxation oscillations of the optical intensity is apparent from the form of (16.5) with $I(t)$ equal to zero. When the drive current is suddenly turned on, damped relaxation oscillations can occur at a frequency ω_R given by

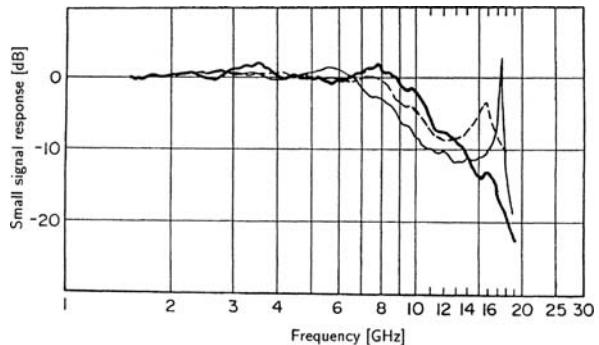
$$\omega_R^2 = \omega_0^2 - (\gamma/2)^2, \quad (16.18)$$

where ω_0 and γ are as defined in Section 16.1.1. In pulsed laser applications, these oscillations take the form of *ringing* of the leading edge of the optical pulse, and can significantly increase the bit error rate. A detailed study of the waveforms of such oscillations and their effect on high data-rate modulation of laser diodes has been performed by Channin et al. [38, 39].

Lau and Yariv [40] have reported that distortion (peaking) of the modulation frequency response curve due to relaxation oscillation resonance can be eliminated by reducing the reflectivity of one of the laser endfaces. Such feedback suppression reduces the photon lifetime, and can extend the flat range of the modulation response curve.

By referring to (16.6), (16.7), (16.10), (16.11) one can see that v_{\max} and the modulation depth both can be increased in three ways – by decreasing the photon lifetime τ_{ph} , by increasing the photon density n_{ph} , and by increasing the optical gain coefficient g . Note that increasing the drive level I_{dc}/I_{th} accomplishes all three of these desired effects. The response curve for a laser designed especially for microwave modulation is shown in Fig. 16.3 [41]. The dark solid curve is the laser response when loosely coupled to an external cavity, and the light solid curve is the response when tightly coupled. This laser has a short cavity length ($\sim 100 \mu\text{m}$) and unequal endface reflectivities to increase the peak photon density and decrease photon lifetime, and it was operated at a high drive level. A window structure was incorporated at the output facet so that a higher photon density could be reached without producing facet damage. The 3 dB bandwidth of this laser was approximately 10 GHz, and response could be peaked at higher frequencies by coupling the laser to an external cavity. The effect of a high drive level on the microwave response of a laser diode can be seen in the data of Fig. 16.4 [42]. The device in that case was a $1.3 \mu\text{m}$ GaInAsP buried-heterostructure laser with a short cavity length of $100 \mu\text{m}$.

Fig. 16.3 Modulation response of a high frequency laser diode [41]



Cooling the device slightly to 20°C improves the frequency response somewhat by increasing the gain coefficient. The observed 18 GHz bandwidth of this laser was determined to be the result of a convolution of an inherent device bandwidth of 22 GHz and a 10 GHz roll-off frequency due to electrical parasitics. To make a very high frequency laser one must make the parasitics small as well as making the resonance frequency high. The constricted-mesa laser [43], shown in Fig. 16.5, has a very low capacity because the bonding pad capacitance is typically only 0.35 pF due to the thick polyimide layer under the bonding pad, and the bonding pad itself is only 50–100 μm wide. The capacitance of the current-confining constricted-mesa

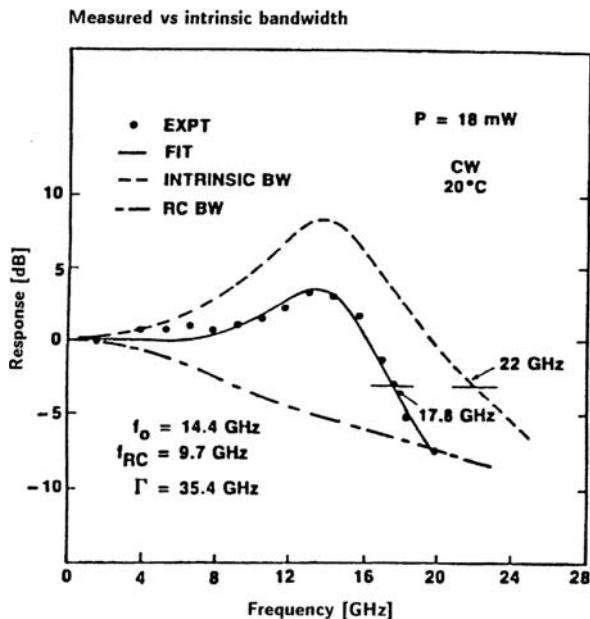
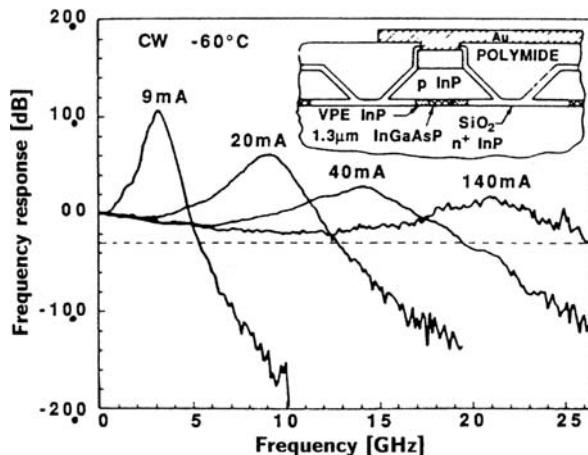


Fig. 16.4 Frequency response of a GaInAsP BH laser [42]

Fig. 16.5 Frequency response of a constricted mesa laser [43]



structure is also very low (0.18 pF). Once again, the strong effect of drive level can be seen in the response data of Fig. 16.5.

Similar results have been obtained by Chen et al. [44] who used an airbridge contact configuration and a narrow mesa to reduce parasitic effects, resulting in a 3 dB bandwidth of 18 GHz in a DFB laser emitting at $1.3 \mu\text{m}$. Their device also incorporated a Multi-Quantum-Well (MQW) structure which further enhances laser performance. (The improvements in laser performance resulting from the inclusion of a MQW structure are described in detail in Section 18.2). Modulation of a MQW DFB laser at $1.55 \mu\text{m}$ wavelength with a bandwidth of 20 GHz has been reported by Lipsanen et al. [45]. In that case, the intrinsic bandwidth of the laser was estimated to be 35 GHz in the absence of parasitics.

The effect of bias current on maximum modulation frequency, which was observed in the laser of Fig. 16.5, is also apparent in the performance of an InGaAs/GaAs/AlGaAs, MQW ridge laser that has been described by Ralston et al. [46]. They observed a modulation bandwidth of 24 GHz at a bias current of 25 mA , which increased to 33 GHz at a bias of 65 mA . Threshold current was approximately 10 mA in these devices. Although the positive effect on modulation bandwidth of increasing bias current is clearly apparent, it should be noted that the current levels in these lasers are still relatively low. This can be attributed to a sophisticated laser geometry employing a short cavity, MQW and ridge-waveguide structure.

When one considers modulation frequencies in the mm-wave range (above 30 GHz) it is best to use a traveling-wave approach. Tromberg et al. [47] have developed a traveling-wave analysis of a general class of semiconductor lasers which includes multisection DFB and DBR lasers as well as gain-coupled DFB lasers. Both intensity modulation and frequency modulation are considered. For a historical review of high-frequency-modulated laser diodes see, for example Suematsu and Arai [48].

16.2.3 Design of Laser Diode Packages for Microwave Modulation

In addition to the four preceding frequency limiting factors inherent in the laser diode chip itself, there are a number of factors associated with the device package and microwave circuit that must be considered as well. Parasitic capacitance and inductance of the laser package are most often the dominant factors that limit modulation frequency when conventional low-frequency laser diodes are used at frequencies above 1 GHz. A diagram of a typical low-frequency laser diode package is shown in Fig. 16.6. It can be seen that the bias lead forms a coaxial capacitor with the threaded stem of the header, which has a capacitance of typically 3 pF. Thus, at a frequency of 10 GHz, a shunt capacitive reactance of only about 5Ω is presented to the microwave signal when it is applied to the bias lead. Because the lead wire from the feedthrough post of the laser chip has an inductance of about 2 nH, corresponding to a series inductive reactance of approximately 140Ω at 10 GHz, it is clear that any attempt to apply a 10 GHz modulating signal to the laser chip via the bias lead will result in virtually all of the signal being shunted to ground by the capacitance.

The unshielded bond wire spanning the gap between the feedthrough post and the diode chip also acts as an *antenna* at frequencies on the order of 10 GHz, permitting significant amounts of microwave energy to be radiated by the laser diode. Not only is this energy lost for optical generation, but it can cause serious RFI (radio frequency interference) by coupling to nearby devices and circuits.

Additional loss of microwave modulating energy occurs in the conventional laser diode package because of rf absorption in the dielectric material used to insulate the bias lead and in the metallic surfaces.

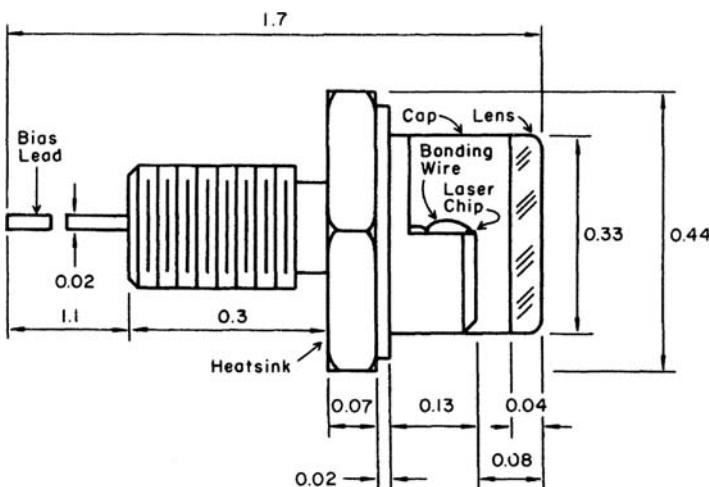


Fig. 16.6 Diagram of a typical commercially available laser diode package. All dimensions are shown in inches

The combined action of all of the microwave loss mechanisms that have been described, coupled with the fact that much of the incident microwave energy is reflected because of the poor impedance match between the diode package and the waveguide or transmission line, leads to the inevitable result that little or no microwave energy actually reaches the laser chip. Obviously, a specially-designed laser diode package is needed for microwave frequency applications.

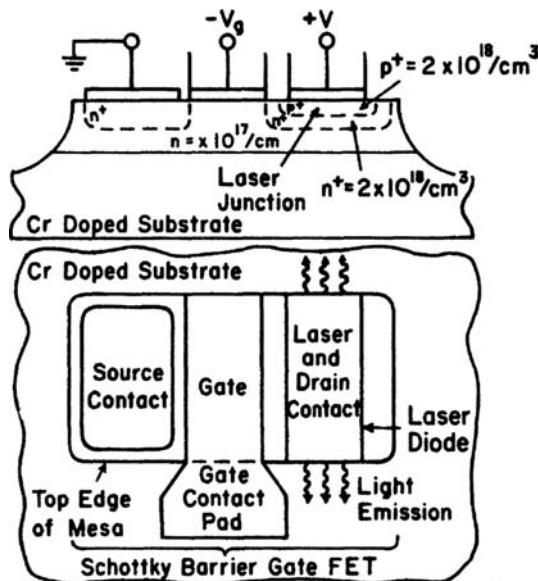
The most effective way of avoiding the deleterious effects of the diode package is to mount the laser diode directly into a microwave strip-line [49]. This type of microwave transmission line, which consists of a relatively thin metallic strip on the surface of a metal-backed dielectric substrate, can be conveniently fabricated on semi-insulating GaAs or InP substrates as well as on ceramic substrates. When designed to the proper dimensions for the frequency band in use, such a strip-line has very low parasitics and very low microwave loss. Commercially available laser diodes intended for use at frequencies above 1 GHz are usually mounted on a heat sink with a ribbon electrical lead designed to properly match a strip-line feed. The fact that microwave strip-lines can be fabricated on semi-insulating GaAs or InP substrates makes the monolithic integration of microwave-modulated laser diodes into optical integrated circuits feasible.

16.3 Monolithically Integrated Direct Modulators

The prospect of monolithically integrating the laser diode with its microwave modulator is attractive, not only because of its obvious applicability to optical integrated circuits, but also because it provides the benefit of reducing the parasitic inductance and capacitance of the interconnecting circuit. An example of a possible monolithic combination of a laser diode with a FET modulator is shown in Fig. 16.7. In this device, the lasing junction is formed as an integral part of the drain region of the field effect transistor, with the two being electrically connected in series. The gate of the FET is controlled by the microwave modulating signal, producing a corresponding modulation of the drain current, and hence, the current through the laser diode. Schottky-barrier-gate FET's formed on high resistivity GaAs substrates, as shown in Fig. 16.7, are known to have cut-off frequencies in the range from 10 to 100 GHz, depending mostly on the gate length [50]. Because of the monolithic integration of the laser junction with the FET drain, interconnection inductance and capacitance are minimized. Hence, improved high frequency modulation response is to be expected.

The fabrication technology required to produce monolithic integration of laser diodes with microwave modulating devices is quite complex, requiring expertise in two fields that are usually dealt with separately. Nevertheless there are a number of examples of such integration being accomplished successfully. Margalit et al. [51] produced a monolithically integrated GaAlAs laser and FET on a Cr-doped GaAs substrate that operated at frequencies above 1 GHz. Ury [52] and Fukuzawa [53] also produced integrated MESFET and laser chips in GaAs. Following these early

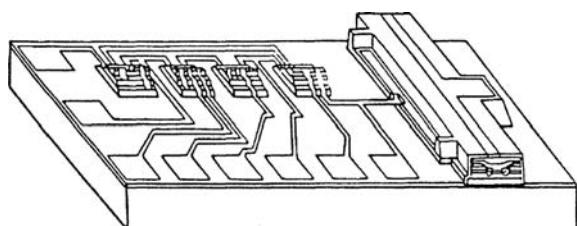
Fig. 16.7 A monolithically integrated laser diode and FET modulator on a GaAs substrate



demonstrations of laser and FET integration, other researchers have gone on to fabricate monolithically integrated transmitters [54–59], receivers [60–65], repeaters [66, 67] and transceivers [68, 69]. These circuits which combine both optical and electronic devices have come to be known as opto-electronic integrated circuits or OEIC's. An example of an OEIC transmitter is shown in Fig. 16.8 [57]. This circuit combines a buried-heterostructure laser with three heterojunction bipolar transistor (HBT) drivers and a FET preamplifier. The HBT's are used as drivers because of their high current capability, while the FET is used as a preamplifier because of its high input impedance. This OEIC is capable of operating at a data rate greater than 21 Gb/S.

A monolithic laser/modulator structure such as those shown in Fig. 16.7 and Fig. 16.8 not only increases the efficiency of microwave coupling, but also provides an opportunity for improved optical coupling, since the laser can be monolithically coupled to a waveguide, as discussed in Chapter 14.

Fig. 16.8 Long wave-length OIC transmitter in GaInAsP. The substrate is semi-insulating InP [57]



Ota et al. [70] have described a monolithically integrated vertical-cavity-surface-emitting laser (VCSEL) with an external cavity on a GaAs substrate. The length of the external cavity was 300 μm . Simulations have shown this combination to have an expected bandwidth of 40 GHz under direct modulation for single-longitudinal-mode operation with out-of-phase coupling of feedback from the external cavity.

16.4 Amplified Laser Modulation

Several factors that limit the maximum modulation frequency of a semiconductor laser are proportional to the level of current passing through the diode. These include frequency shift during a pulse (chirping) due to change in the index of refraction caused by injected carriers, heating effects and delays caused by parasitic capacitance and inductance. One way to combat these factors is to operate the laser diode at a low power level, with a relatively low drive current, and to subsequently amplify the laser output to the desired power level for transmission. Any of the optical amplifiers discussed in Chapter 13 can be used, but generally a SOA is the most compatible. If the laser and the amplifier are made in the same semiconductor material system they can be conveniently integrated. For example, Verdiell et al. [71] have combined a DBR laser with a SOA in AlGaAs/InGaAs/GaAs to make a monolithically integrated master oscillator/power amplifier (MOPA), as shown in Fig. 16.8.

This device is capable of producing an output of 3 W cw at a wavelength of 980 nm, and is intended for use in satellite optical communications systems. The laser is modulated at a low power to avoid the effects mentioned previously and the output is then amplified to the high power level. Since the stimulated emission in the SOA is proportional to the number of signal photons entering it, the modulation signal is reproduced at the higher power level. A DBR laser was used to ensure single mode operation and wavelength stability. The modulation bandwidth under direct modulation is 3 GHz. As an alternative to directly modulating the current through the laser diode, it is also possible to incorporate an electroabsorption modulator section within the device which increases the bandwidth to 5 GHz.

16.5 Direct Modulation of Quantum Dot Lasers

Quantum dot (QD) lasers, such as those described in Chapters 15 and 22, have been shown to be capable of being directly modulated at frequencies up to 20 GHz. For example, Gerschütz et al. [72] have made a laser with a 1.3 μm InGaAs/GaAs QD active region. Their device featured a special multisectional coupled-cavity-injection-grating design that allowed interaction between the lasing mode and a second mode used as catalyst. This approach produced a high self resonance frequency of the device that permitted the 20 GHz modulation bandwidth. There were 3 sections, gain (450 μm), DFB grating (1000 μm), and phase, with a total cavity length

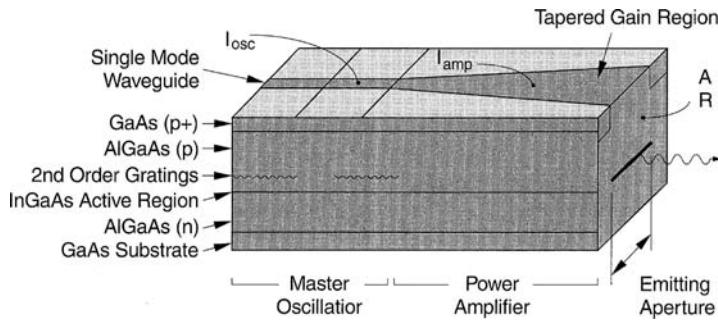


Fig. 16.9 Monolithic MOPA laser [71]

of 2.4 mm. The modulating signal was impressed on the current flow in the gain section.

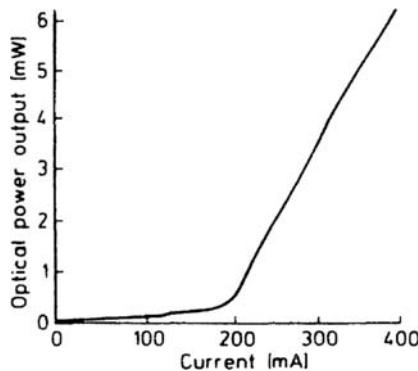
A DFB laser with InGaAs QDs in a GaAs active layer was also reported by Kunz et al. [73]. Their device, which emitted at 1.3 μm , was a 1 μm -thick ridge waveguide structure. The small-signal modulation cutoff frequency was 7.4 GHz. The laser was operated in the injection-locked mode to produce pulse trains with repetition rates from 5 to 50 GHz. The minimum pulse width obtained was 3 ps.

16.6 Future Prospects for Microwave Modulation of Laser Diodes

In this chapter, a variety of techniques for directly modulating semiconductor lasers at frequencies into the mm-wave range have been reviewed. Some impressive results have been obtained, but much work remains to be done before the ultimate limits imposed by carrier and photon lifetimes are reached. New OEIC technology offers the promise of driver circuits with negligible parasitics. Bandwidths as high as 40 GHz have been reached with OEICs [74] and nanotechnology has enabled the fabrication of new laser structures, such as the QD laser, that also can be directly modulated at frequencies up to 40 GHz. However, this work is still mostly confined to the laboratory. Commercially, lasers directly modulated up to 10 GHz can be obtained from a number of different suppliers, but higher frequency capability is rare. There is little doubt though that the 40 GHz devices will be commercialized soon.

Problems

- 16.1 We wish to modulate the light output of a semiconductor laser having the optical power-current characteristic shown below by directly varying the input current. Sketch the intensity (or optical power density) waveform of the light



output for a bias current of 300 mA and an applied current signal which is sinusoidal and has a peak to peak current variation of 200 mA.

- 16.2 Repeat problem 16.1 for a bias current of 150 mA, all other conditions remaining the same.
- 16.3 A cw double heterojunction GaAs-GaAlAs laser diode is to be directly current modulated by a sinusoidal ac signal current of 100 mA peak-to-peak. What must be the minimum dc bias current to the laser to ensure that the laser output linearly follows the input signal, i.e., the output optical signal will be sinusoidal, following the input signal current waveform. Relevant characteristics of the laser diode are as follows:
- (1) Emitted wavelength: 9000 Å.
 - (2) Half power points of the emission peak for spontaneous emission have been measured for this material at room temperature to be at 9200 and 8800 Å.
 - (3) Index of refraction: 3.3.
 - (4) Thickness of active layer 1 μm.
 - (5) Internal quantum efficiency: 0.8.
 - (6) Average absorption coefficient: 10 cm^{-1} .
 - (7) Length between cleaved end faces: 1 mm.
 - (8) Reflectivity of end faces: 0.4.
 - (9) Cross-sectional area normal to current flow: 10^{-3} cm^2 .
- 16.4 Given a laser with threshold current $I_t = 100 \text{ mA}$ and peak current $I_p = 300 \text{ mA}$, what is the maximum modulation depth achievable if the dc bias current is $I_{dc} = 130 \text{ mA}$? Assume that the laser L - I curve is linear above threshold and that, at threshold, the output power $P_t = 0.05 P_p$.
- 16.5 A SH laser diode is to be used for high-speed communications. It must be operated with no dc bias, and in a pulsed mode with no more than a 5% duty cycle, and a pulse width of 0.8 ns. The threshold and peak currents are 100 mA and 7 A, respectively. The spontaneous electron lifetime in the device is 60 ns.

- (a) What factor(s) limit the maximum modulation rate?
 - (b) What is the maximum pulse repetition rate that can be used?
- 16.6 In general, what factors limit the maximum modulation frequency that can be attained by directly modulating the drive current of a semiconductor diode laser?
- 16.7 (a) Why is direct modulation of the drive current of a semiconductor laser not used at modulation frequencies above 10 GHz.
- (b) Describe a method that could be used to modulate a semiconductor laser at 50 GHz.
- 16.8 A directly modulated laser with a threshold current of 1 mA is used to occasionally transmit digital data over an optical fiber from a remote location to a central monitoring center. To save on power consumption the laser is operated at zero bias when not transmitting. If the peak pulse current of the laser is 5 mA and its carrier lifetime for spontaneous emission is 300 ns, what is the delay time before the first pulse is transmitted after drive current is applied?
- 16.9 In a dual cavity laser emitting at a wavelength of 950 nm, one cavity has a length of 300 μm and an effective index of 3.45 while the second cavity has a length of 320 μm and an effective index of 3.47. What is the mode spacing of the coupled cavity?

References

1. G.J. Lasher: Solid State Electron. **7**, 707 (1964)
2. T.L. Paoli, J.E. Ripper: IEEE Proc. **58**, 1457 (1970)
3. T. Ikegami, Y. Suematsu: Electron. Commun. (Jpn.) B **51**, 51 (1968)
4. T.P. Lee, R.M. Derosier: IEEE Proc. **62**, 1176 (1974)
5. K. Konnerth, C. Lanza: Appl. Phys. Lett. **4**, 120 (1964)
6. M. Ross: IEEE Trans. AES-**3**, 324 (1967)
7. G.E. Fenner: Pulse width modulated laser. US Patent no. 3,478,280 (Nov, 1969)
8. L.A. D'Asaro, J.M. Cherlow, T.L. Paoli: IEEE J. **QE-4**, 164 (1968)
9. J.E. Ripper, T.L. Paoli: Appl. Phys. Lett. **15**, 203 (1969)
10. T.L. Paoli, J.E. Ripper: IEEE J. **QE-6**, 335 (1970)
11. G.E. Fenner: Appl. Phys. Lett. **5**, 198 (1964)
12. W.T. Tsang, N.A. Olsson, R.A. Logan: Appl. Phys. Lett. **42**, 650 (1983)
13. L.A. Coldren, D.I. Miller, K. Iga, A. Rentschler: Appl. Phys. Lett. **38**, 315 (1981)
14. W.T. Tang, N.A. Olsson, R.A. Logan: IEEE J. **QE-19**, 1621 (1983)
15. W. Streifer, D. Yevick, T.L. Paoli, R.D. Burnham: IEEE J. **QE-20**, 754 (1984)
16. J.E. Ripper: IEEE J. **QE-6**, 129 (1970)
17. H. Reick: Solid State Electron. **8**, 83 (1965)
18. D. Kleinman: Bell Syst. Tech. J. **43**, 1505 (1964)
19. B. Goldstein, R. Wigand: IEEE Proc. **53**, 195 (1965)
20. R. Myers, P. Pershan: J. Appl. Phys. **36**, 22 (1965)
21. T. Nakano, T. Oku: Jpn. J. Appl. Phys. **6**, 1212 (1967)
22. T. Ikegami, Y. Suematsu: IEEE Proc. **55**, 122 (1967)
23. T. Ikegami, Y. Suematsu: Electron. Commun. (Jpn.) B **51**, 51 (1968)
24. J. Nishizawa: IEEE J. **QE-4**, 143 (1968)

25. T. Ikegami, Y. Suematsu: IEEE J. **QE-4**, 148 (1968)
26. S. Takamiga, F. Kitasawa, J. Nishizawa: IEEE Proc. **56**, 135 (1968)
27. Zh.I. Alferov: Sov. Phys. – Semicond. **3**, 1170 (1970)
28. M.B. Panish, I. Hayashi, S. Sumski: Appl. Phys. Lett. **17**, 109 (1970)
29. M. Lakshminarayana, R.G. Hunsperger, L. Partain: Electron. Lett. **14**, 640 (1978)
30. M. Maeda, K. Nagano, I. Ikushima, M. Tanaka, K. Saito, R. Ito: 3rd Europ. Conf. Opt. Commun., NTG Fachberichte **59**, 120 (1977)
31. H. Yania, M. Yano, T. Kamiya: IEEE J. **QE-11**, 519 (1975)
32. J. Carroll, J. Farrington: Electron. Lett. **9**, 166 (1973)
33. P. Russer, S. Schultz: Arch. Elektr. Übertrag. **27**, 193 (1973)
34. T. Ozeki, T. Ito: IEEE J. **QE-9**, 388 (1973)
35. A.J. Seeds, J.R. Forrest: Electron. Lett. **14**, 829 (1978)
36. H.W. Yen: OSA/IEEE Conf. on Laser Engineering and Applications, Washington, DC (1979) Digest p. 9D
37. F. Mengel, V. Ostoich: IEEE J. **QE-13**, 359 (1977)
38. D.J. Channin, M. Ettenberg, H. Kressel: J. Appl. Phys. **50**, 6700 (1979)
39. D.J. Channin: SPIE Proc. **224**, 128 (1980)
40. K.Y. Lau, A. Yariv: Appl. Phys. Lett. **40**, 452 (1982)
41. K.Y. Lau, A. Yariv: Appl. Phys. Lett. **46**, 326 (1985)
42. R. Olshansky, V. Lanzisera, C. Bsu, W. Powazinik, R.B. Lauer: Appl. Phys. Lett. **49**, 128 (1986)
43. J.E. Bowers: Solid State Electron. **30**, 1 (1987)
44. T. Chen, P. Chen, J. Ungar, N. Bar-Chain: Electron. Lett. **30**, 1055 (1994)
45. H. Lipsanen, D. Coblenz, R. Logan, R. Yadvish, P. Morton, H. Temkin: IEEE Photon. Tech. Lett. **4**, 673 (1992)
46. J. Ralston, S. Weisser, K. Eisele, R. Sah, E. Larkins, J. Rosenzweig, J. Fleissner, K. Bender: IEEE Photon. Tech. Lett. **6**, 1076 (1994)
47. B. Tromberg, H. Lassen, H. Olesen: IEEE J. **QE-30**, 939 (1994)
48. Y. Suematsu, S. Arai: Single-mode semiconductor lasers for long-wavelength optical fiber communications and dynamics of semiconductor lasers. IEEE J. Selected Topics Quantum Elect. **6**, 1436 (2000)
49. P.A. Rizzi: *Microwave Engineering Passive Circuits* (Prentice-Hall, Englewood Cliffs, NJ 1988) pp. 248–299
50. S. Dods, M. Ogura, M. Watanabe: IEEE J. **QE-29**, 2631 (1993)
51. S. Margalit, N. Bar-Chaim, I. Ury, D. Wilt, M. Yust, A. Yariv: Monolithic integration of optical and electronic devices on semi-insulating GaAs substrates. OSA/IEEE Topical Meeting on Integrated and Guided-Wave Optics, Incline Village, NV (1980)
52. I. Ury, S. Margalit, M. Yust, A. Yariv: Appl. Phys. Lett. **34**, 430 (1979)
53. T. Fukuzawa, M. Nakamura, M. Hirao, T. Kuroda, J. Umeda: Appl. Phys. Lett. **36**, 181 (1980)
54. H. Nakano, S. Yamashita, T. Tanaka, N. Hirao, N. Naeda: IEEE J. LT-**4**, 574 (1986)
55. K. Dretting, W. Idler, P. Wiedermann: Electron. Lett. **29**, 2195 (1993)
56. P. Woolnough, P. Birdsall, P. O’Sullivan, A. Cockburn, M. Harlow: Electron. Lett. **29**, 1388 (1993)
57. N. Suzuki, H. Furuyama, Y. Hirayama, M. Morinaga, E. Eguchi, M. Kushibe, M. Funamizu, M. Nakamura: Electron. Lett. **24**, 467 (1988)
58. K. Kudo, K. Yashiki, T. Sasaki, Y. Yokoyama, K. Hamamoto, T. Morimoto, M. Yamaguchi: 1.55 μm wavelength-selectable microarray DFB-LD’s with monolithically integrated MMI combiner, SOA, and EA-modulator. IEEE Photonics Tech. Lett. **12**, 242 (2000)
59. S. Menezo, A. Rigny, A. Talneau, F. Delorme, S. Grosmaire, H. Nakajima, E. Vergnol, F. Alexandre, F. Gaborit: Design, realization, and characterization of a ten-wavelength monolithic source for WDM applications integrating DBR lasers with a PHASAR. IEEE J. Selected Topics Quant. Electron. **6**, 185 (2000)

60. J. Wang, C. Shih, N. Chang, J. Middleton, P. Apostolakis, M. Feng: IEEE Photon. Tech. Lett. **5**, 316 (1993)
61. J. Wang, C. Shih, W. Chang, J. Middleton, P. Apostolakis, M. Feng: IEEE MTT-S Int'l Symp. on Circuits and Systems, Atlanta, GA (1993) Digest Vol. 2, p. 1047
62. D. Trommer, Unterborsch, F. Reier: 5th Int'l Conf. on InP and Related Materials, Paris (1993) Proc. p. 251
63. C. Shih, D. Barlage, J. Wang, M. Feng: IEEE MTT-S Int'l Microwave Symp., San Diego, CA (1994) Digest, Vol. 3, p. 1379
64. M. Bitter, R. Bauknecht, W. Hunziker H. Melchior: Monolithic InGaAs-InP p-i-n/HBT 40-Gb/s optical receiver module. IEEE Photonics Tech. Lett. **12**, 74 (2000)
65. R. Li, J.D. Schaub, S.M. Csutak, J.C. Campbell: A high-speed monolithic silicon photoreceiver fabricated on SOI. IEEE Photonics Tech. Lett. **12**, 1046 (2000)
66. M. Yust, N. Bar-Chaim, S. Izapanah, S. Margalit, I. Ury, D. Wilt, A. Yariv: Appl. Phys. Lett. **35**, 796 (1979)
67. S. Yamashita, D. Matsumoto: Waveform reshaping based on injection locking of a distributed-feedback semiconductor laser. IEEE Photonics Tech. Lett. **12**, 1388 (2000)
68. K. Jackson, E. Flint, M. Cina, D. Lacey, Y. Kwark, J. Trewella, T. Caulfield, P. Buchmann, Ch. Harder, P. Vettiger: IEEE J. LT-**12**, 1185 (1994)
69. H. Nakajima, A. Leroy, J. Charil, D. Robein: Versatile in-line transceiver chip operating in two full-duplex modes at 1.3 and 1.55 μm . IEEE Photonics Tech. Lett. **12**, 202 (2000)
70. T Ota, T. Ochida, T. Kondo, F. Koyama: Enhanced modulation bandwidth of surface-emitting laser with external optical feedback, IEICE Electron. Express **1**, 368 (2004)
71. J.M. Verdelli, R.J. Lang, K. Dzurko, S. O'Brien, J. Osinsky, D.F. Welch, D.R. Scifres: Monolithically integrated high-speed, high-power, diffraction limited semiconductor sources for space telecommunications. Proc. SPIE **2684**, 108 (1996)
72. F. Gerschütz1, M. Fischer, J. Koeth, I. Krestnikov, A. Kovsh, C. Schilling, W. Kaiser, S. Höfling, A. Forchel: 1.3 μm quantum dot laser in coupled-cavity-injection-grating design with bandwidth of 20 GHz under direct modulation, Opt. Express, **16**, 5596 (2008)
73. M. Kuntz, G. Fiol, M Lämmelin, D. Bimberg, M.G. Thompson, K.T. Tan, C. Marinelli, A. Wonfor, R. Sellin, R.V. Penty, I.H. White, V.M. Ustinov, A.E. Zhukov, Yu M. Shernyakov3, A.R. Kovsh, N.N. Ledentsov, C. Schubert, V. Marembert: Direct modulation and mode locking of 1.3 μm quantum dot lasers, IOP New J. Phys. **6**, 181 (2004)
74. T. Ohno, K. Sato, S. Fukushima, Y. Doi, Y. Matsuoka: Application of DBR mode-locked lasers in millimeter-wave fiber-radio system. IEEE J. Lightwave Tech. **18**, 44 (2000)

Supplementary Reading on Modulation of Laser Diodes

- G. Arnold, P. Russer, K. Petermann: Modulation of laser diodes, in *Semiconductor Devices for Optical Communication*, H. Kressel 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982) Chap. 7
- J.K. Butler (ed.): *Semiconductor Injection Lasers* (IEEE Press, New York 1980) pp. 332–389
- H.C. Casey Jr., M.B. Panish: *Heterostructure Lasers* (Academic, New York 1978) pp. 258–264

Chapter 17

Integrated Optical Detectors

Detectors for use in integrated-optic applications must have high sensitivity, short response time, large quantum efficiency and low power consumption [1]. In this chapter, a number of different detector structures having these performance characteristics are discussed.

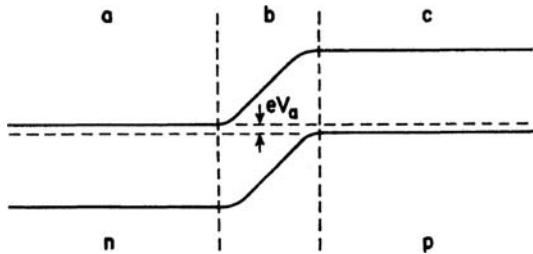
17.1 Depletion Layer Photodiodes

The most common type of semiconductor optical detector, used in both integrated optic and discrete device applications, is the depletion-layer photo-diode. The depletion-layer photodiode is essentially a reverse-biased semiconductor diode in which reverse current is modulated by the electron-hole pairs produced in or near the depletion layer by the absorption of photons of light. The diode is generally operated in the *photodiode* mode, with relatively large bias voltage, rather than in *photovoltaic* mode, in which the diode itself is the electrical generator and no bias voltage is applied [2].

17.1.1 Conventional Discrete Photodiodes

The simplest type of depletion layer photodiode is the p-n junction diode. The energy band diagram for such a device, with reverse bias voltage V_a applied is shown in Fig. 17.1. The total current of the depletion layer photodiode consists of two components: a drift component originating from carriers generated in region (b) and a diffusion component originating in regions (a) and (c). Holes and electrons generated in region (b) are separated by the reverse bias field, with holes being swept into the p-region (c) and electrons being swept into the n-region (a). Holes generated in the n-region or electrons generated in the p-region have a certain probability of diffusion to the edge of the depletion region (b), at which point they are swept across by the field. Majority carriers, electrons in (a) or holes in (c) are held in their respective regions by the reverse bias voltage, and are not swept across the depletion layer.

Fig. 17.1 Energy band diagram for a p-n junction diode under application of a reverse bias voltage V_a



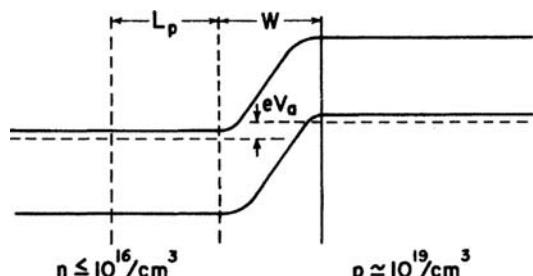
In order to minimize series resistance in a practical photodiode while still maintaining maximum depletion width, usually one region is much more heavily doped than the other. In that case, the depletion layer forms almost entirely on the more lightly doped side of the junction, as shown in Fig. 17.2. Such a device is called a high-low abrupt junction. In GaAs and its ternary and quaternary alloys, electron mobility is generally much larger than hole mobility. Thus, the p-region is usually made thinner and much more heavily doped than the n-region, so that the device will be formed mostly in n-type material, and the p-region then serves essentially just as a contact layer.

For a device with the high-low junction geometry indicated in Fig. 17.2, it can be shown that the total current density J_{tot} is given by [3]

$$J_{\text{tot}} = q\varphi_0 \left(1 - \frac{e^{-\alpha W}}{(1 + \alpha L_p)} \right) + qp_{n0} \frac{D_p}{L_p} \quad (17.1)$$

where φ_0 is the total photon flux in photons/cm²s, W is the width of the depletion layer, q is the magnitude of the electronic charge, α is the optical interband absorption coefficient, L_p is the diffusion length for holes, D_p is the diffusion constant for holes, and P_{n0} is the equilibrium hole density. The last term of (17.1) represents the reverse *leakage* current (or *dark* current), which results from thermally generated holes in the n-material. This explains why that term is not proportional to the photon flux φ_0 . The first term of (17.1) gives the photocurrent, which is proportional

Fig. 17.2 Energy band diagram for a p^+ -n (high-low) junction diode under application of a reverse bias voltage V_a



to φ_0 , and includes current from both the drift of carriers generated within the depletion layer and the diffusion and drift of holes generated within a diffusion length L_p of the depletion layer edge. The quantum efficiency η_q of the detector, or the number of carriers generated per incident photon, is given by

$$\eta_q = 1 - \frac{e^{-\alpha W}}{(1 + \alpha L_p)}, \quad (17.2)$$

which can have any value from zero to one. It should be noted that (17.1) and (17.2) are based on the tacit assumption that scattering loss and free carrier absorption are negligibly small. The effect of these loss mechanisms on the quantum efficiency, when they are not negligible, is discussed in Section 17.1.3.

It can be seen from (17.2) that, in order to maximize η_q , it is desirable to make the products αW and αL_p as large as possible. When αW and αL_p are large enough so that η_q is approximately equal to one, the diode current is then essentially proportional to φ_0 , because the dark current is usually negligibly small.

If the interband absorption coefficient α is too small compared to W and L_p , many of the incident photons will pass completely through the active layers of the diode into the substrate, as shown in Fig. 17.3. Only those photons absorbed within the depletion layer, of thickness W , have maximum effectiveness in carrier generation. Photons absorbed at depths up to a diffusion length L_p from the depletion layer edge are somewhat effective in generating photo-carriers, in that holes can diffuse into the depletion layer. Photons that penetrate to a depth greater than $(W + L_p)$ before being absorbed are essentially lost to the photo-generation process because they have such a very low statistical probability of producing a hole that can reach the depletion layer and be swept across. Within the semiconductor, the photon flux $\varphi(x)$ falls off exponentially with increasing depth x from the surface, as shown in Fig. 17.4. Thus, if α is not large enough, many photons will penetrate too deeply before being absorbed, thus producing carriers that (on average) will recombine before diffusing far enough to reach the depletion layer.

Interband absorption is a strong function of wavelength in a semiconductor. The absorption coefficient α usually exhibits a special response curve that rises sharply at the absorption-edge (band-edge) wavelength and then saturates at a wavelength that is

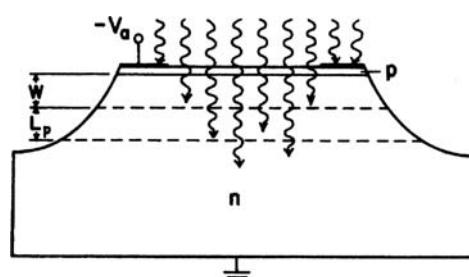
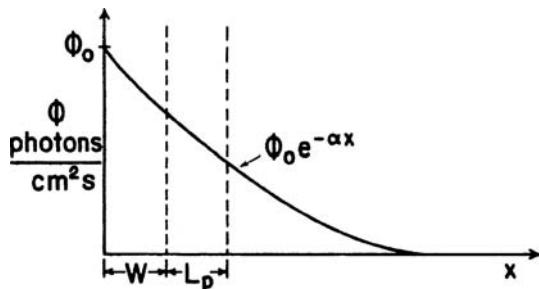


Fig. 17.3 Diagram of a conventional mesa-geometry photodiode with p⁺-n doping profile showing photon penetration

Fig. 17.4 Optical absorption versus depth from the surface in a conventional mesa photodiode



slightly shorter than the bandgap wavelength, increasing slowly for yet shorter wavelengths. Thus, it is impossible to design a diode with an ideal W for all wavelengths. For wavelengths near the absorption edge, the long-wavelength response of a diode is limited by excess penetration of photons into the substrate, as shown in Figs. 17.3 and 17.4; its short wavelength response can be limited by too strong an absorption of photons in the p^+ layer near the surface, where recombination probability is large.

Aside from the reduction of quantum efficiency that results from poor matching of α , W and L_p , there are some other limitations to depletion layer photodiode performance that are also important. Since W is usually relatively small (in the range from 0.1 to 1.0 μm), junction capacitance can limit high-frequency response through the familiar R-C time constant. Also, the time required for carriers to diffuse from depths between W and $(W + L_p)$ can limit the high frequency response of a conventional photodiode. The waveguide depletion layer photodiode, which is discussed in the next section, significantly mitigates many of these problems of the conventional photodiode.

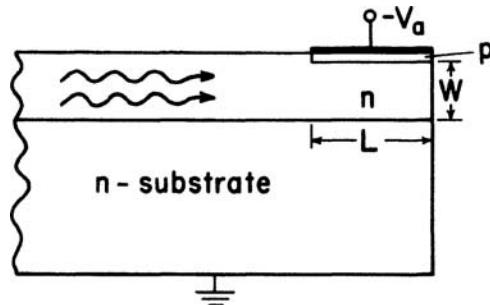
17.1.2 Waveguide Photodiodes

If the basic depletion layer photodiode is incorporated into a waveguide structure, as shown in Fig. 17.5, a number of improvements in performance are realized. In this case, the light is incident transversely on the active volume of the detector, rather than being normal to the junction plane. The diode photocurrent density is then given by

$$J = q\varphi_0 (1 - e^{-\alpha L}), \quad (17.3)$$

where L is the length of the detector in the direction of light propagation. Since W and L are two independent parameters, the carrier concentration within the detector volume and the bias voltage V_a can be chosen so that the depletion layer thickness W is equal to the thickness of the waveguide, while L can be made as long as necessary to make $\alpha L \gg 1$. Thus 100% quantum efficiency can be obtained for any value of α , by merely adjusting the length L . For example, for a material with the relatively small value of $\alpha = 30 \text{ cm}^{-1}$, a length of $L = 3 \text{ mm}$ would give $\eta_q = 0.99988$. (Again, it

Fig. 17.5 Diagram of a waveguide detector



has been tacitly assumed in (17.3) that scattering loss and free-carrier absorption are negligible.)

Because a waveguide detector can be formed in a narrow channel waveguide, the capacitance can be very small, even if L is relatively large. For example, for a material with a relative dielectric constant $\epsilon = 12$, such as GaAs, a 3 mm long detector formed in a 3 μm wide channel waveguide has a capacitance of only 0.32 pF. This capacitance is about a factor of ten less than that of a typical conventional mesa photodiode. Hence, the high frequency response can be expected to be correspondingly improved. Experimentally demonstrated bandwidth of 5 GHz and quantum efficiency of 83% have been obtained with waveguide detectors on GaAs substrate material [4], and InGaAs waveguide photodetectors on InP substrates also have exhibited a 5 GHz bandwidth for light in the wavelength range of 1.3–1.6 μm [5]. Computer simulation of waveguide photodetectors in AlGaInAs–GaInAs, designed for ultrawide-band operation at 60 and 100 GHz, predict internal quantum efficiencies as high as 94% and 75%, respectively, at 1.55 μm wavelength [6].

Because all of the incident photons are absorbed directly within the depletion layer of a waveguide photodetector, not only is η_q improved, but also the time delay associated with the diffusion of carriers is eliminated. This result is a further improvement in high frequency response.

Due to the many improvements in performance inherent in the transverse structure of the waveguide detector, as compared to the axial geometry of the conventional mesa photodiode, waveguide detectors should be considered for use in discrete-device applications, as well as in optical integrated circuits. At the present time, waveguide detectors are not commercially available as discrete devices. However, they can be fabricated with relative ease in many laboratories. Hence availability should not long be a problem.

17.1.3 Effects of Scattering and Free-Carrier Absorption

The relations given by (17.1), (17.2) and (17.3) neglected the effects of free-carrier absorption and photon scattering on the quantum efficiency of the detector. Because

both of these mechanisms result in the loss of photons without the generation of any new carriers, they tend to reduce quantum efficiency. In many cases they can be neglected, and (17.1), (17.2) and (17.3) will give accurate predictions. However, when the free carrier absorption coefficient α_{FC} and the scattering loss coefficient α_s are not negligible as compared to the interband absorption coefficient α_{IB} , a more sophisticated expression for η_q is required. Such an expression can be derived as follows. The photon flux at any point located a distance x from the surface of the detector on which the photons are first incident is assumed to have the form given by

$$\varphi(x) = \varphi_0 e^{-\alpha x}, \quad (17.4)$$

where in general the loss coefficient α is given by

$$\alpha = \alpha_{\text{IB}} + \alpha_{\text{FC}} + \alpha_s. \quad (17.5)$$

The hole-electron pair generation rate $G(x)$ is given by

$$G(x) = \alpha_{\text{IB}} \varphi_0 e^{-\alpha x}, \quad (17.6)$$

since only α_{IB} results in carrier generation. Thus the photocurrent density is given by

$$J = q \int_0^L G(x) dx \quad (17.7)$$

or

$$J = q \varphi_0 \frac{\alpha_{\text{IB}}}{\alpha_{\text{IB}} + \alpha_{\text{FC}} + \alpha_s} \left(1 - e^{(\alpha_{\text{IB}} + \alpha_{\text{FC}} + \alpha_s)L} \right). \quad (17.8)$$

Comparing (17.8) with (17.3), it is obvious that the effect of additional losses due to scattering and free-carrier absorption is to reduce the quantum efficiency by a factor of $\alpha_{\text{IB}}/\alpha$, even when L is large enough to maximize η_q .

If α_s and α_{FC} are small compared to α_{IB} , as is generally true, (17.8) reduces to (17.3). However, if the waveguide is inhomogeneous or is unusually rough, or if the detector volume is heavily doped so that α_s and α_{FC} are not negligible, then (17.8) must be used.

17.2 Specialized Photodiode Structures

There are two very useful photodiode structures that can be fabricated in either a waveguiding or conventional, nonwaveguiding form. These are the Schottky-barrier photodiode and the avalanche photodiode.

17.2.1 Schottky-Barrier Photodiode

The Schottky-barrier photodiode is simply a depletion layer photodiode in which the p-n junction is replaced by a metal-semiconductor rectifying (blocking) contact. For example, if the p-type layers in the devices of Figs. 17.3 and 17.5 were replaced by a metal that forms a rectifying contact to the semiconductor, Schottky-barrier photodiodes would result. The photocurrent would still be given by (17.1) and (17.3), and the devices would have essentially the same performance characteristics as their p⁺-n junction counterparts. The energy band diagrams for a Schottky-barrier diode, under zero bias and under reverse bias, are given in Fig. 17.6. It can be seen that the depletion region extends into the n-type material just as in the case of a p⁺-n junction. The barrier height φ_B depends on the particular metal-semiconductor combination that is used. Typical values for φ_B are about 1 V.

In conventional mesa devices, a thin, optically transparent Schottky-barrier contact is often used (rather than a p⁺-n junction) to enhance short-wavelength response, by eliminating the strong absorption of these higher energy photons that occurs in the p⁺ layer. In a waveguide photodiode, a Schottky-barrier contact is not needed for improved short-wavelength response because the photons enter the active volume transversely. However, ease of fabrication often makes the Schottky-barrier photodiode the best choice in integrated applications. For example, almost any metal (except for silver) produces a rectifying Schottky-barrier when evaporated onto GaAs or GaAlAs at room temperature. Gold, aluminum or platinum are often used. Transparent conductive oxides such as indium Tin oxide (ITO) and cadmium Tin oxide (CTO) can also be used to eliminate the photon masking effect of the contacts and thereby improve the quantum efficiency, as described in Section 17.2.4. Photore sist masking is adequate to define the lateral dimensions during evaporation, and no careful control of time and temperature is required, as in the case of diffusion of a shallow p⁺ layer.

A detailed discussion of the properties of Schottky-barrier diodes is beyond the scope of this text so that the interested reader should refer to the information available elsewhere [7].

17.2.2 Avalanche Photodiodes

The gain of a depletion layer photodiode (i.e. the quantum efficiency), of either the p-n junction or Schottky-barrier type, can be at most equal to unity, under normal

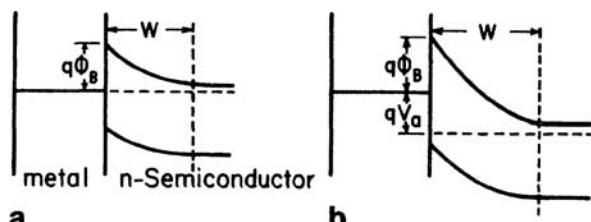
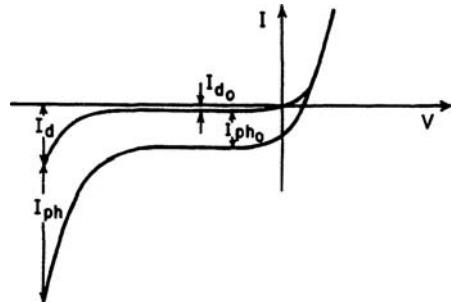


Fig. 17.6 a,b Energy band diagram for a Schottky-barrier diode: **a** zero bias; **b** reverse biased with voltage V_a

Fig. 17.7 Response curves for an avalanche photodiode



conditions of reverse bias. However, if the device is biased precisely at the point of avalanche breakdown, carrier multiplication due to impact ionization can result in substantial gain in terms of increase in the carrier to photon ratio. In fact, avalanche gains as high as 10^4 are not uncommon. Typical current-voltage characteristics for an avalanche photodiode are shown in Fig. 17.7. The upper curve is for darkened conditions, while the lower one shows the effects of illumination. For relatively low reverse bias voltages, the diode exhibits a saturated dark current I_{d0} and a saturated photocurrent I_{ph0} . However, when biased at the point of avalanche breakdown, carrier multiplication results in increased dark current I_d , as well as increased photocurrent I_{ph} . It is possible to define a photomultiplication factor M_{ph} , given by

$$M_{ph} \equiv \frac{I_{ph}}{I_{ph0}}, \quad (17.9)$$

and a multiplication factor M , given by

$$M \equiv \frac{I_{ph} + I_d}{I_{ph0} + I_{d0}}, \quad (17.10)$$

An exact equation for the current-voltage curve is difficult to obtain in the region of bias in which avalanche breakdown occurs. However, Miller [8] has represented the functional form of the photomultiplication factor by the expression

$$M_{ph} = \frac{1}{1 - (V_a/V_b)^n}. \quad (17.11)$$

where V_b is the breakdown voltage, and n is an empirically determined exponent depending on the wavelength of light, doping concentration, and, of course, the semiconductor material from which the diode is fabricated. For the case of large photocurrent $I_{ph0} \gg I_{d0}$ Melchior and Lynch [9] have shown that the multiplication factor is given by

$$M = \frac{1}{1 - \left(\frac{V_a - IR}{V_b} \right)^n}, \quad (17.12)$$

where I is the total current, given by

$$I = I_d + I_{ph}, \quad (17.13)$$

R being the series resistance of the diode (including space-charge resistance if significant). The derivation of (17.12) assumes that $IR \ll V_b$. For the case of I_{d0} and I_d being negligibly small compared to I_{ph0} and I_{ph} , it can be shown that the maximum attainable multiplication factor is given by [9]

$$M \cong M_{ph} \cong \sqrt{\frac{V_b}{n I_{ph0} R}}. \quad (17.14)$$

Avalanche photodiodes are very useful detectors, not only because they are capable of high gain, but also because they can be operated at frequencies as high as 35 GHz [10, 11]. However, not every p-n junction or Schottky-barrier diode can be operated in the avalanche multiplication mode, biased near avalanche breakdown. For example, the field required to produce avalanche breakdown in GaAs is approximately 4×10^5 V/cm. Hence, for a typical deplection width of $5 \mu\text{m}$. V_b equals 120 V. Most GaAs diodes will breakdown at much voltages lower due to other mechanisms, such as edge breakdown or microplasma generation at localized defects, thus never reaching the avalanche breakdown condition. In order to fabricate an avalanche photo-diode, extreme care must be taken, beginning with a dislocation free substrate wafer of semiconductor material. Generally a guard ring structure [7, p. 203] must be employed to prevent edge breakdown.

Avalanche photodiodes are highly-stressed devices. Hence, reliability is a question of prime concern. Increasing leakage current due to poor surface passivation or the generation of internal defects during high current pulse operation can lead to degradation of performance as the devices age. Nevertheless, when diodes are carefully fabricated and are hermetically sealed into adequate packages, mean time to failure as high as 10^5 h at 170°C has been observed [1, p. 80], which projects to about 10^9 h at room temperature.

17.2.3 p-i-n Photodiodes

In Section 17.1.1. it was pointed out that conventional photodiodes must be designed so as to have a large αW product in order to maximize η_q ; but one doesn't have complete control over either the depletion width W , which depends on dopant concentrations, or the absorption coefficient, which depends mostly on the bandgap. In the p-i-n photodiode, a very lightly doped “intrinsic” layer is formed between the p and n sides of the diode. This layer generally has a carrier concentration of less than

$10^{14}/\text{cm}^3$, but it is compensated by a balance of p- and n-type dopants rather than being truly intrinsic. Because of the low carrier concentration, the depletion layer in a p-i-n diode extends completely through the i layer so that the total thickness of the active layer is the sum of the i-layer thickness W_i and the depletion width on the lightly doped (n) side of the junction. Thus the device designer can adjust the total depletion width to produce a large αW product by varying the thickness of the i-layer. The presence of the relatively thick i-layer also reduces the junction capacitance and increases the R-C cutoff frequency of the diode, p-i-n photo-diodes are widely used as detectors in optical systems because of their high quantum efficiency (responsivity) and wide bandwidth. For example, Kato et al. [12] have reported a waveguide p-i-n photodiode operating at $1.55 \mu\text{m}$ wavelength with a quantum efficiency of 50% and a 3 dB bandwidth of 75 GHZ.

17.2.4 Metal-Semiconductor-Metal Photodiodes

Metal-semiconductor-Metal (MSM) photodiodes are surface-oriented devices that feature interdigitated, finger-like, Schottky barrier contacts formed on the surface of a thin semiconducting layer on a semi-insulating substrate. A typical MSM photodiode structure is shown in Fig. 17.8. Carriers generated by the absorption of photons in regions of the semiconducting layer between the contacts are swept by the fringing electric field and collected by the contacts. Holes are collected by the cathode and electrons by the anode. Spacing of the contact fingers must be less than the diffusion length of the carriers in order to produce a high collection efficiency.

Because the contact fingers are very narrow and closely spaced ($\sim 1 \mu\text{m}$) the capacitance is relatively low and the transit times of carriers are short. Hence, wide-bandwidth operation is possible. Hsiang et al. have made Si MSM diodes with $0.2 \mu\text{m}$ width and spacing with a full-width-half-maximum pulse response of 3.7 ps, corresponding to a 3 dB bandwidth at 110 GHz [13]. The Schottky electrodes of the MSM photodiode are essentially identical to the gate metalization of field-effect transistors, which facilitates their monolithic integration with FET's. For example, Mactaggart et al. [14] have made a fully-integrated 400 Mb/s burst-mode data OEIC receiver for application as a phased-array antenna controller. Approximately 350 source-coupled FET logic gates are present on the GaAs chip, along with a 780 nm wavelength MSM photodiode. MSM photodiodes have also been integrated

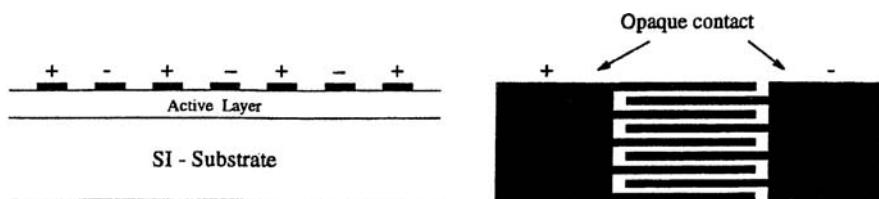


Fig. 17.8 Conventional MSM Photodetector

monolithically with *High-Electron-Mobility* field effect Transistors (HEMT's) to produce OEIC receivers with bandwidths larger than 14 GHz [15, 16]. HEMT's have also been integrated with p-i-n photodiodes to produce OEIC receivers with a bandwidth of 42 GHz [17].

The most significant disadvantage of MSM detectors is their inherent low responsivity because the metallization for the surface electrodes shadows the active light-collecting region. However, this problem can be mitigated by using a transparent conducting material for the contact electrodes. Gao et al. [18] have fabricated InGaAs MSM photodiodes with transparent Cadmium Tin Oxide (CTO) electrodes. The responsivity of these devices to $1.3\text{ }\mu\text{m}$ light was 0.49 A/W , as compared to 0.28 A/W for identical control samples with conventional Ti/Au electrodes.

Another approach to improve the overall responsivity of an m-s-m photodiode is to monolithically integrate it with an amplifier. For example, Cha et al. [19] have integrated an m-s-m photodiode with a high-electron-mobility transistor (HEMT) on an InP substrate, with an InGaAsP buffer layer ($\lambda_g = 1.3\text{ }\mu\text{m}$). They measured a responsivity of 0.7 A/W at a wavelength of $\lambda=1.3\text{ }\mu\text{m}$, and the $1.5 \times 100\text{ }\mu\text{m}^2$ gate HEMT had f_t and f_{max} of 18.7 and 47 GHz, respectively.

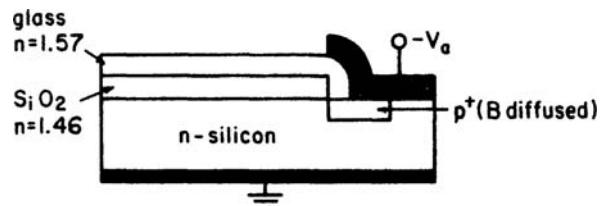
17.3 Techniques for Modifying Spectral Response

The fundamental problem of wavelength incompatibility, which was encountered previously in regard to the design and fabrication of monolithic laser/waveguide structures in Chapter 14, is also very significant with respect to waveguide detectors. An ideal waveguide should have minimal absorption at the wavelength being used. However a detector depends on interband absorption for carrier generation. Hence, if a detector is monolithically coupled to a waveguide, some means must be provided for increasing the absorption of the photons transmitted by the waveguide within the detector volume. A number of different techniques have proven effective in this regard.

17.3.1 Hybrid Structures

One of the most direct approaches to obtaining wavelength compatibility is to use a hybrid structure, in which a detector diode, formed in a relatively narrow bandgap material, is coupled to a waveguide fabricated in wider bandgap material. The two materials are chosen so that photons of the desired wavelength are transmitted freely by the waveguide, but are strongly absorbed within the detector material. An example of this type of hybrid waveguide/detector is the glass on silicon structure that was demonstrated by Ostrowsky et al. [20], as shown in Fig. 17.9. The diode was formed by boron diffusion to a depth of about $1\text{ }\mu\text{m}$ into an n-type, $5\text{ }\Omega\text{cm}$ silicon substrate. A $1\text{ }\mu\text{m}$ thick layer of thermally grown SiO_2 was used as a diffusion mask. The glass waveguide was then sputter-deposited and silver paint electrodes were

Fig. 17.9 Hybrid waveguide detector featuring a glass waveguide coupled to a silicon photodiode [20]



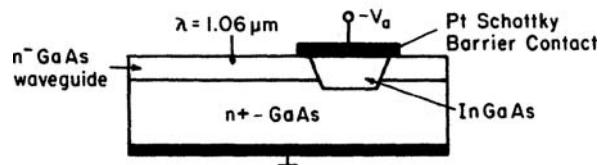
added as shown. Total guide loss was measured to be $0.8 \text{ dB/cm} \pm 10\%$ for light of 6328 \AA wavelength. The efficiency of coupling between the waveguide and the detector was 80%. However, because the light enters the diode in the direction normal to the junction plane rather than parallel to it, this particular waveguide detector geometry does not have many of the advantages described in Section 17.1.2. Nevertheless, good high frequency response can be expected. These diffused diodes had a capacitance of only $3 \times 10^{-9} \text{ F/cm}^2$ when reverse biased with V_a equal to 10 V. Thus a detector diode of approximately $10 \mu\text{m}$ radius, used in conjunction with a 50Ω load resistance, would have an RC time constant of about 15 ps, implying that modulation of frequencies in excess of 10 GHz could be detected. Another example of hybrid coupling is provided by Koike et al. [21], who have coupled a flip-chip Ge photodiode to a dielectric waveguide with a coupling efficiency of 22%.

While hybrid detectors offer the possibility of choosing the waveguide and detector materials for optimum absorption characteristics, better coupling efficiency can be obtained with monolithic fabrication techniques. Monolithically fabricated waveguide detectors also have the advantage that light enters the device in the plane of the junction rather than normal to it.

17.3.2 Heteroepitaxial Growth

The most popular method of monolithically integrating a waveguide and detector is to use heteroepitaxial growth of a relatively narrow bandgap semiconductor at the location where a detector is desired. An example of this approach is given by the InGaAs detector that has been integrated with a GaAs waveguide by Stillman et al. [22], as shown in Fig. 17.10. In $\text{In}_x\text{Ga}_{(1-x)}\text{As}$, the bandgap can be adjusted to produce strong absorption of light at wavelengths in the range from 0.9 to $3.5 \mu\text{m}$

Fig. 17.10 Monolithically integrated InGaAs detector in a GaAs waveguide [22]



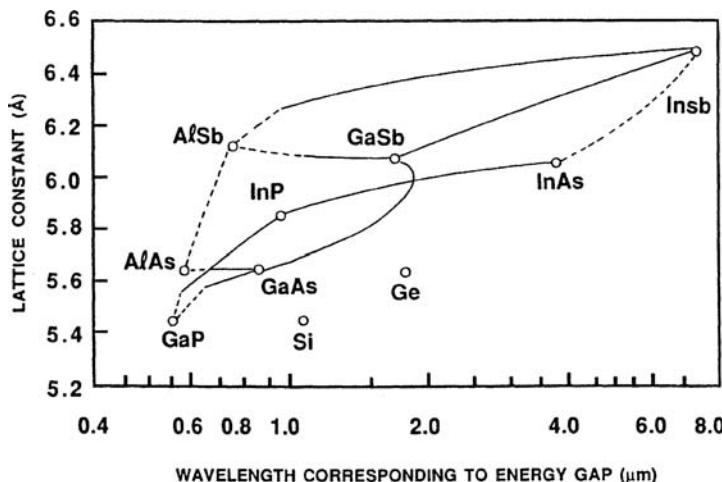
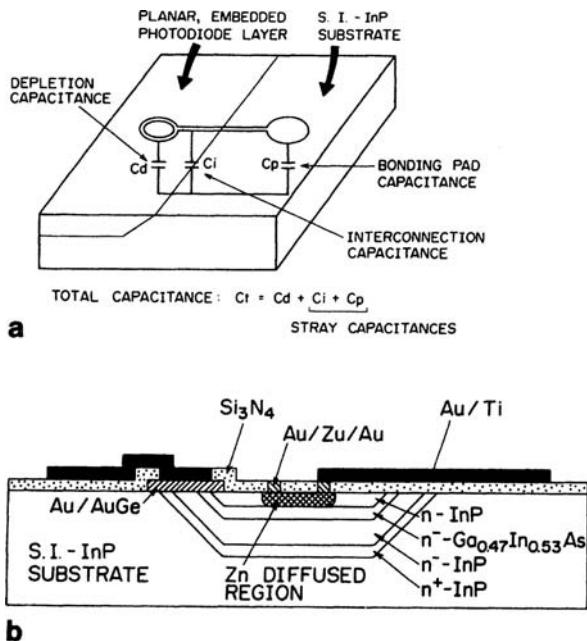


Fig. 17.11 Dependence of absorption edge wavelength and lattice constant on composition for selected III-V alloys [23]

by changing the atomic fraction x of indium (Fig. 17.11). The monolithic waveguide detector structure shown in Fig. 17.12 combines an epitaxially grown carrier-concentration-reduction type waveguide with a platinum Schottky barrier detector. A 6000 Å thick layer of pyrolytically deposited SiO_2 was used as a mask to etch a 125 μm diameter well into the 5–20 μm thick waveguide, and then grow the $\text{In}_x\text{Ga}_{(1-x)}\text{As}$ detector material. A quantum efficiency of 60% was measured for this detector at a wavelength of 1.06 μm, for low bias voltages. The loss in the waveguide was less than 1 dB/cm. At bias voltage greater than about 40 V, avalanche multiplication was observed, with multiplication factors as high as 50. Optimum performance was obtained at a wavelength of 1.06 μm for an In concentration of $x = 0.2$. The fact that quantum efficiency in this device did not approach 100% more closely was most likely caused by less than optimum depletion width in the Schottky-barrier diode. The carrier concentration in the waveguide must be very carefully controlled in order to make W equal to the waveguide thickness. Defect centers associated with stress at the GaAs-GaInAs interface may have also played a role in reducing η_q .

In general, the III-V compound semiconductors and their associated ternary (and quaternary) alloys offer the device designer a wide range of bandgaps, and corresponding absorption edge wavelengths. The relationships between bandgaps, absorption edge wavelengths, and lattice constants are shown in Fig. 17.11, as given by Kimura and Daikoku [23]. Dotted portions of the curves correspond to ranges of composition for which the bandgap is indirect. Direct bandgap materials generally have interband absorption coefficients greater than 10^4 cm^{-1} for wavelengths shorter than the absorption edge, while α may be several orders of magnitude less in indirect gap materials. Nevertheless effective detectors can be made in indirect

Fig. 17.12 a A planar embedded GaInAs p-i-n photodiode [24], **b** cross-section of the planar embedded photodiode [24]



bandgap materials, especially when the waveguide detector geometry is used, so that the length L can be adjusted to compensate for small α .

The most popular material for the fabrication of detectors in the 1.0 to $1.6\text{ }\mu\text{m}$ wavelength range is GaInAsP. Figure 17.12 shows a planar embedded GaInAs p-i-n photodiode on an InP substrate [24]. The device was fabricated by means of preferential ion beam etching and vapor phase epitaxy. Since the bonding pads lie mostly over semi-insulating substrate material, stray capacitances have been reduced to the point at which total device capacitance is only 0.08 pF for a diameter of $20\text{ }\mu\text{m}$ of photosensitive area. As a result, this device has a cutoff frequency of 14 GHz, limited by the carrier transit time. An integrated waveguide p-i-n photodetector coupled to a ridge waveguide on InP is shown in Fig. 17.13 [25]. The device was fabricated by means of a metal organic vapor phase epitaxial (MOVPE) regrowth technique. The detector can be used in the $1.0\text{--}1.6\text{ }\mu\text{m}$ wavelength range; and has a 3 dB bandwidth of 1.5 GHz with a pulse response (full-width-at-half-maximum) of 80 ps at $1.3\text{ }\mu\text{m}$ wavelength. The waveguides had an average propagation loss of 3 dB/cm and 95% of the guided light was coupled into the photodetector.

Another approach to integrating a GaInAs detector and an InP waveguide is shown in Fig. 17.14 [26]. In this case an InGaAs detector was grown on top of the waveguide. Since InGaAs has a larger index of refraction than InP, light is coupled out of the waveguide up into the detector.

Another type of coupling structure that can be heteroepitaxially grown to couple a narrower-bandgap photodetector to a wider-bandgap waveguide is the

Fig. 17.13 A p-i-n photodiode integrated with a ridge waveguide [25]

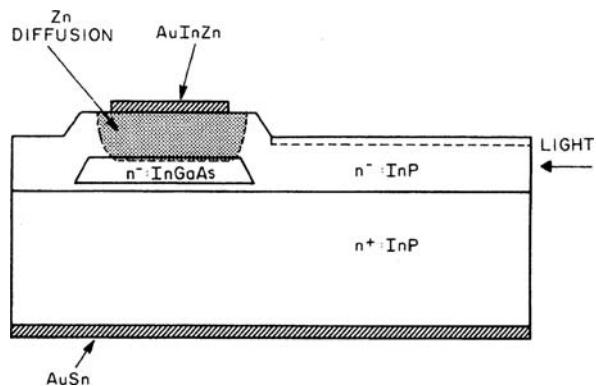
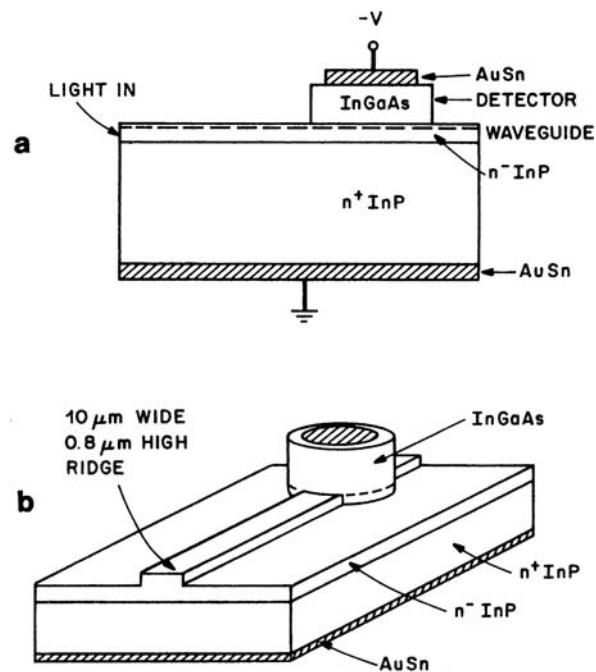


Fig. 17.14 An InGaAs detector grown on top of a ridge waveguide [26]



evanescent-wave or optical-tunneling coupler. This type of coupler depends on the overlap of the evanescent tails of the modes traveling in two adjacent waveguides. (See Figs. 8.1 and 8.2.) Demiguel et al. [27] have made an evanescently coupled photodiode that utilizes a short planar multimode waveguide. They report a responsivity greater than 1 A/W, polarization dependence less than 0.5 dB, 48-GHz bandwidth, and 11-mA saturation current.

17.3.3 Proton Bombardment

In Chapter 4, proton bombardment was described as a method for producing optical waveguides in a semiconductor by generating carrier-trapping defect centers that resulted in reduced carrier concentration and increased index of refraction. In that case, the waveguides were always sufficiently annealed after proton bombardment to remove the optical absorption associated with the trapping centers. However, one of the mechanisms responsible for this absorption is the excitation of carriers out of the traps, freeing them to contribute to the flow of photocurrent. Thus, a photodiode can be fabricated by forming a Schottky-barrier junction over the implanted region, as shown in Fig. 17.15 [28], (A shallow p⁺-n junction could also be used). Photocurrent flows when the junction is reverse biased, because carriers liberated by photo-excitation within the depletion layer of the diode are swept across it by the field. Since a substantial number of the trapping centers have energy levels lying within the forbidden gap, the effective bandgap of the semiconductor is decreased, so that photons having less-than-bandgap energy can be absorbed and take part in the carrier generation process. Thus a photon-bombarded photodiode, made in a given carrier generation process. Thus a proton-bombarded photodiode, made in a given semiconductor, can be responsive to photons that would ordinarily not be absorbed in the material.

For example, Stoll et al. [28] have made a detector in GaAs that is sensitive to 1.15 μm wavelength radiation. The optical waveguide structure consisted of a 3.5 μm thick n-type epitaxial layer (S-doped, $n \cong 10^6 \text{ cm}^{-3}$) grown on a degenerately doped n-type substrate ($n \cong 1.25 \times 10^{18} \text{ cm}^{-3}$). Prior to proton implantation the optical attenuation at 1.15 μm was measured to be 1.3 cm⁻¹, but after implantation with a dose of $2 \times 10^{15} \text{ cm}^{-2}$ 300 keV protons in the region where a detector was desired, α increased to over 300 cm⁻¹. A partial annealing of damage at 500°C for 30 min was performed to reduce α to 15 cm⁻¹ in order to allow some optical transmission through the entire length of the implanted region. Then, an Al Schottky-barrier contact was evaporated on top of the implanted region to complete the device. The relative photoresponse of the proton implanted detector as a function

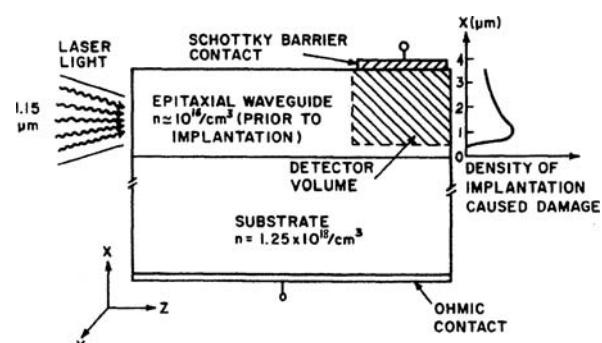
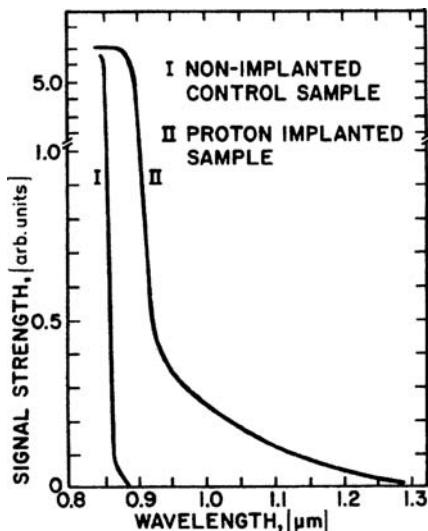


Fig. 17.15 Diagram of a proton-implanted optical detector [28]

Fig. 17.16 Photoresponse of a proton implanted detector in GaAs



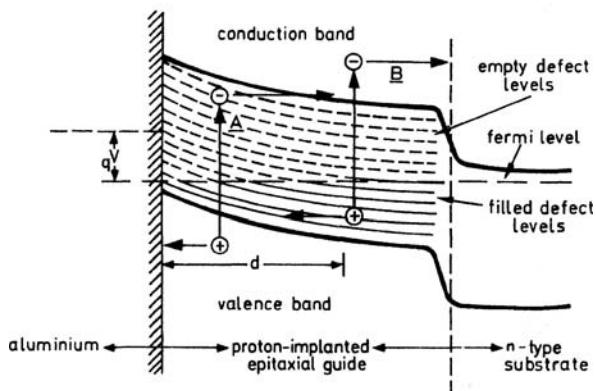
of wavelength is shown in Fig. 17.16, along with the corresponding curve for a similar detector formed in unimplanted GaAs. Only negligible response occurs in the unimplanted GaAs at wavelengths longer than 9000 Å, but a substantial absorption tail is observed for the proton bombarded detector, extending to wavelengths as long as 1.3 μm .

It must be remembered that, even though α may be relatively small at the longer wavelengths, total absorption over the length of the detector can be quite large. For example, quantum efficiencies as high as 17% have been measured at 1.15 μm wavelength for devices of this type that were only 0.25 mm long. Calorimetric measurements made on proton-bombarded GaAs at a wavelength of 1.06 μm indicate that essentially all of the bombardment-induced optical attenuation can be attributed to absorption rather than diffuse scattering [29]. Calculations indicate that interband (carrier producing) absorptive transitions make up about 60% of the total absorption [29]. Thus quantum efficiency as high as 60% should be possible with this type of detector.

The principle of operation of the proton-bombarded detector is similar to that of conventional depletion layer photodiodes. Application of a reverse bias to the Schottky barrier produces a depletion layer which, if bias is sufficient and carrier concentration is low enough, extends completely across the high resistivity waveguiding layer to the lower resistivity substrate. Photoexcitation of trapped carriers from bombardment-produced defect levels generates free carriers that are swept out of the depletion layer by the field, as shown in Fig. 17.17, thereby causing photocurrent to flow in the external circuit.

Proton bombarded detectors thusfar have been fabricated in GaAs and GaAlAs. However, the same bombardment-generated defect trapping mechanism is known

Fig. 17.17 Energy band diagram of a proton-bombarded detector under reverse bias and illumination



to exist in GaP [30], ZnTe [31] and CdTe [32]. Hence, it seems likely that proton bombarded detectors can be made in these materials, and perhaps others as well.

The response time of proton-bombarded detectors is relatively fast for a device which depends on trapping and detrapping. Measurements made in GaAs with a Q-switched Nd:YAG laser at $1.06 \mu\text{m}$ wavelength indicate response times less than 200 ns, which was the rise-time of the laser pulse.

A proton implanted detector has been employed by Carenco and Menigaux [33] to make an optically bistable device, in GaAs, as shown in Fig. 17.18. The device consists of a rib-waveguide dual-channel directional coupler with a Schottky-barrier proton implanted detector in one of its output arms. The photocurrent output of the detector is amplified by an external electronic circuit and fed back to the control electrodes of the directional coupler switch to provide bistable response to an optical input signal. Operating at a wavelength of $1.06 \mu\text{m}$, the bistable device switches in about 1 ns, the optical switching energy required being less than 1 nJ. The quantum

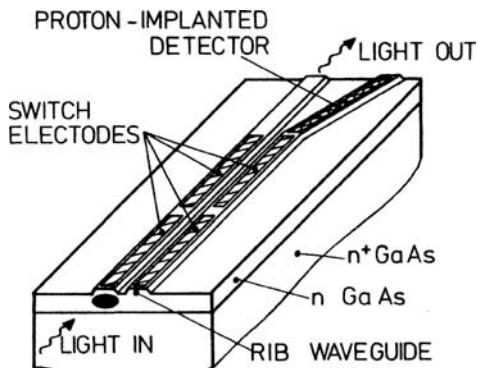


Fig. 17.18 An optically bistable device [33]

efficiency of the proton implanted detector was measured to be 23% at $1.06 \mu\text{m}$ and 13% at $1.15 \mu\text{m}$. The estimated bandwidth of the detector was 1.2 GHz.

17.3.4 Electro-Absorption

One additional method for producing the required shift of the absorption edge to longer wavelength in a monolithic waveguide detector is electro-absorption, or the Franz-Keldysh effect. When a semiconductor diode is reverse biased, a strong electric field is established within the depletion region. This electric field causes the absorption edge to shift to longer wavelengths, as shown in Fig. 17.19. Curve A shows the normal unbiased absorption edge for n-type GaAs with a carrier concentration of $3 \times 10^{16} \text{ cm}^{-3}$. Curve B is a calculated Franz-Keldysh-shifted absorption edge for an applied field of $1.35 \times 10^5 \text{ V/cm}$, which corresponds to 50 V reverse bias across a resulting depletion width of $3.7 \mu\text{m}$. At a wavelength of 9000 \AA this shift corresponds to an increase in α from 25 to 10^4 cm^{-1} – hardly a negligible effect!

The Franz-Keldysh effect has been well known for many years [34]. However it has only been applied to detector design fairly recently. The physical basis for the Franz-Keldysh effect can be understood from the simplified energy and bending model diagrammed in Fig. 17.20. In this diagram x represents distance from the metallurgical junction plane. In the region far from the junction where there is no

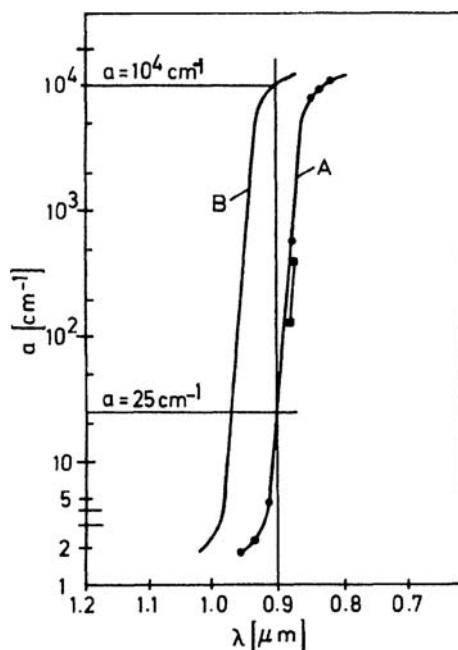
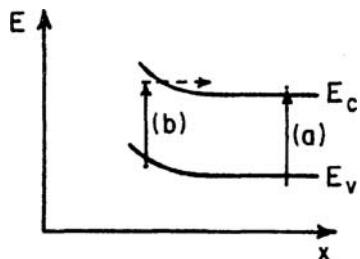


Fig. 17.19 Shift of the absorption edge of GaAs due to the Franz-Keldysh effect.
(A) Zero-bias condition; (B) Reverse bias applied to produce a field of $1.35 \times 10^5 \text{ V/cm}$

Fig. 17.20 Energy band diagram illustrating the Franz-Keldysh effect. The band bending on the n-side of a p⁺-n junction (or a Schottky barrier junction) is shown for conditions of strong reverse bias



electric field, photons must have at least the bandgap energy ($E_c - E_v$) to produce an electronic transition as in (a). However, within the depletion region where field is strong, a transition as in (b) can occur when a photon of less-than bandgap energy lifts an electron part way into the conduction band, followed by tunneling of the electron through the barrier into a conduction band state. The states at the conduction band edge are, in effect, broadened into the gap so as to produce a change in effective bandgap ΔE , which is given by [34]

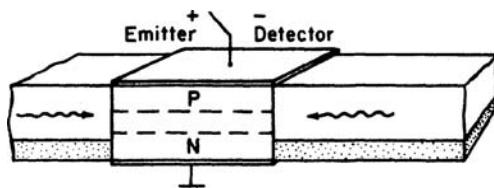
$$\Delta E = \frac{3}{2} (m^*)^{-1/3} (q\hbar\varepsilon)^{2/3}, \quad (17.15)$$

where m^* is the effective mass of the carrier, q is the magnitude of the electronic charge, and ... is the electric field strength.

The Franz-Keldysh effect greatly improves the sensitivity of a detector operating at a wavelength near its absorption edge. GaAs waveguide detectors operating at a wavelength of 1.06 μm have been demonstrated by Nichols et al. [35].

Perhaps the greatest advantage of electro-absorption detectors is that they can be electrically switched from a low absorption state to a high absorption state by merely increasing the reverse bias voltage. This makes it possible to make emitters and detectors in the same semiconductor material that are wavelength compatible. An example of a device making use of this principle is the emitter/detector terminal shown in Fig. 17.21 [36]. This device performs the dual function of light emitter, when forward biased, and light detector, when strongly reverse biased. Fabricated in series with a waveguide structure, as shown, it can act as a send/receive tap on an optical transmission line. Because of the large change in α produced by the Franz-Keldysh effect, operation can be very efficient. For example, consider the case of

Fig. 17.21 An integrated-optic emitter detector terminal employing the Franz-Keldysh effect



a p⁺-n junction diode in n-type GaAs with carrier concentration equal to 3×10^{16} cm⁻³, as before. Application of 50 V reverse bias changes α from 25 to 10⁴ cm⁻¹ at a wavelength of 9000 Å. Thus, when forward biased, the diode emits 9000 Å light into the waveguide. When reverse biased with $V_a = 50$ V, the diode need have a length of only 10⁻³ μm in order to absorb 99.9% of incident 9000 Å light. When the diode is on standby at zero bias, α is just 25 cm⁻¹. Hence, for a typical laser length of 200 μm, the insertion loss is only 2 dB. Such emitter/detector devices may prove to be very useful in systems employing waveguide transmission lines because they greatly simplify coupling problems, as compared to those encountered when using separate emitters and detectors.

Continuing work on light emitting and detecting diode (LEAD) devices has produced a laser/detector in GaAlAs with a differential quantum efficiency of 30.9% in the emission mode, and a responsivity of 0.43 A/W in the detection mode [37, 38]. This responsivity at the operating wavelength of 905 nm is notably close to the 0.5 A/W which can be obtained with a silicon p-i-n diode such as is normally used. Laser/detector diodes have been used in a full duplex (simultaneous bidirectional) single optical fiber 16 km transmission system operating at a bit rate of 40 Mb/s [39]. In this system the laser/detector diodes function simultaneously as a local oscillator, a mixer and a transmitter. Distributed feedback LEAD devices have been reported by Sakano et al. [40]. In their InGaAsP OIC, shown in Fig. 17.22, DFB LEAD devices are used alternatively as lasers, detectors or amplifiers, combined with a carrier injection type switch. Thus, this single QIC chip can function in variety of ways, such as a modulated or switched optical source, a photodiode-monitored source, or an amplified source. The lasers emit at 1.55 μm and are of the constricted-mesa buried heterostructure type. Typical threshold currents and differential quantum efficiencies are 30 mA and 0.08 mW/mA, respectively. When operated as detectors, the quantum efficiency of the diodes was 0.9% in the photovoltaic (zero-bias) mode. Optical amplification gain of up to 17.7 dB has been realized in these diodes [41].

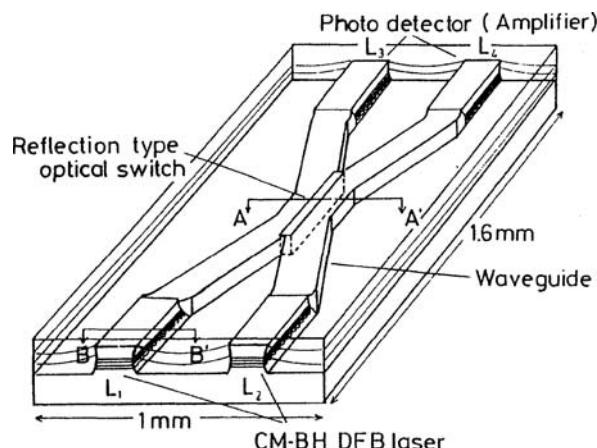


Fig. 17.22 An InGaAsP OIC featuring, DFB LEAD devices [40]

In addition to the electroabsorption effect that results from the Franz-Keldysh effect, which been described in this section and previously also in Section 9.3.4, there is another type of electroabsorption that occurs only in quantum well structures. This latter type of electroabsorption, which depends on the quantum-confined Stark effect (QCSE) is explained in Chapter 18. Both the Franz-Keldysh and QCSE types of electroabsorption can be used in photodetectors, as well as in modulators like those described in Chapter 9. Shin et al. [42] have used both types to make an integrated photodiode/mixer for frequency conversion of radio frequency modulation signals. In that application the QCSE device had greater conversion gain.

17.4 Factors Limiting Performance of Integrated Detectors

In the design of an integrated optical detector, there are a number of mechanisms that can limit performance in various ways. Not all of these are important in every application. However, the designer (or user) should be aware of the limitations associated with different device types and geometries.

17.4.1 High Frequency Cutoff

A number of the factors that can limit high frequency response have been discussed in Section 17.1. These are summarized in Table 17.1, along with some additional frequency-limiting effects. Because of the small area of waveguide photodetectors of the type shown in Fig. 17.5, the RC time constant, which most often limits the response of conventional diodes, can be made small enough to allow frequencies of operation well above 10 GHz, as discussed in Section 17.1.2. In this case, other potentially limiting effects must be considered.

The drift time of carriers across the depletion layer can be minimized by designing the device so that the field in the depletion layer is high enough that carriers travel at the scattering limited velocity. For example, in GaAs the scattering limited velocity of 1×10^7 cm/s is reached in electric field strengths greater than about 2×10^4 V/cm. Hence, the transit time across a typical 3 μm wide depletion layer can be made as small as 3×10^{-11} s. However, it is important that the detector be designed so that the depletion layer extends entirely through the waveguide to the substrate, so that all carriers are generated within the depletion layer itself. If deple-

Table 17.1 Factors limiting high frequency response of a depletion layer photodiode

RC time constant due to bulk series resistance and junction capacitance
Carrier diffusion time from regions outside of the depletion layer
Carrier lifetime and diffusion length
Capacitance and inductance of the package
Carrier drift time across the depletion layer
Carrier trapping in deep levels

tion is incomplete, carriers generated in undepleted material must diffuse relatively slowly into the depletion region before being collected, a process that occurs over roughly a minority carrier lifetime (about 10^{-8} s in lightly doped GaAs). A short optical pulse will therefore appear to have a long *diffusion tail* when detected.

Carrier trapping in deep levels can also cause a *tail* on the detected waveforms of a short pulse, since detrapping times can be relatively long. Deep level traps are usually associated with defects present in the semiconductor crystal lattice. Hence, special care should be taken in materials selection and device fabrication to minimize the number of defects.

When waveguide detectors are properly designed and are appropriately fed with a microwave transmission line, bandwidths greater than 60 GHz [43].

17.4.2 Linearity

A depletion layer photodiode, when reverse biased by more than a couple of volts, operates in the photodiode (or photoconductive) mode. It then functions as a current source, with its current being proportional to the input optical power, up to power levels of typically 1 mW. Hence, in most applications it is a highly linear device. At higher power levels, saturation occurs when the concentration of photo-generated carriers is so large that the field in the depletion layer is reduced by space-charge effects. This field reduction is particularly important in high frequency applications, because it may then result in carriers traveling at less than the saturation-limited velocity.

17.4.3 Noise

The effects of noise in waveguide detectors are essentially the same as they are in conventional photodiodes. The major noise components are thermal noise, arising in bulk resistances of the device, shot noise, associated with nonuniformities of current flow, such as carrier generation and recombination, and background noise, due to photons that are not part of the optical signal entering the detector [44].

From a relatively simple model developed by DiDomenico and Svelto [45], it can be shown that the signal-to-noise ratio in a depletion layer photodiode due to the effects of thermal and shot noise is given by [46]

$$(S/N)_{\text{power}} = \frac{\eta_q}{4B} M^2 \varphi_0 A \left(1 + \frac{2KT}{q} \frac{(\omega RC)^2}{RI_s} \right)^{-1} \quad (17.16)$$

where η_q is the quantum efficiency, B is the bandwidth, M is the modulation index (in the case of an intensity modulated light beam), φ_0 is the incident photon flux density, A is the area of the input face, R is the diode bulk resistance, ω is the modulation frequency of the optical signal, I_s is the reverse saturation (dark) current, K is Boltzmann's constant and C is the capacitance. In the case of an avalanche

photodiode, there are additional noise sources associated with the statistical nature of the avalanche process that are not considered in (17.16) [1, pp. 72–77].

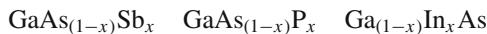
The waveguide detector has an inherent advantage over conventional detectors in regard to background noise, because the waveguide acts as a filter to eliminate much of the background light. Close matching to the signal wavelength also reduces background noise. For example, if light of 8500 Å wavelength from a GaAlAs emitter is detected by a Si detector, background photons with wavelengths shorter than about 1.2 μm will also be detected. However if, instead, a GaAs detector is used, those photons with wavelengths in the range from 1.2 to 0.9 μm will not contribute to background noise [47].

Problems

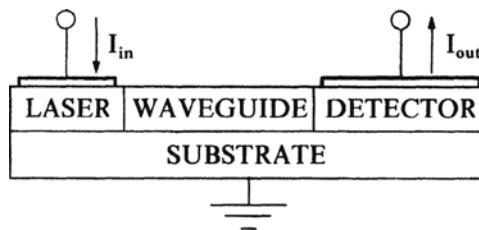
- 17.1 We wish to use a photodiode as a detector for a signal of 9000 Å wavelength.
 Which would be the best choice of material for the photodiode, a semiconductor of bandgap = 0.5 eV, bandgap = 2 eV, or bandgap = 1 eV? Why? (Assume all three are direct gap and are equivalent in impurity content, etc.)
- 17.2 To improve the signal-to-noise ratio of the diode in Problem 17.1, we wish to use a semiconductor low-pass filter which has the following absorption properties at room temperature:
 for 9000 Å radiation $\alpha = 0.2 \text{ cm}^{-1}$
 for 7000 Å radiation $\alpha = 10^3 \text{ cm}^{-1}$.
 How thick must the filter be to attenuate 7000 Å background noise by a factor of 10^4 ?
 By what factor is the signal (at 9000 Å) attenuated by a filter of this thickness?
 Neglect reflection at the surfaces.
- 17.3 If the minimum useful photocurrent of the diode in Problem 17.1 is 1 μA (peak pulse value), what is the minimum signal light intensity (peak pulse value) which must fall on the detector? Assume an internal yield or quantum efficiency $\eta_q = 0.8$ and sensitive area = 10 mm².
- 17.4 Below is a list of semiconductor materials and their bandgaps.

	$E_g[\text{eV}]$
Si	1.1
GaAs	1.4
GaSb	0.81
GaP	2.3
InAs	0.36

- (a) Based solely on bandgap energy, which of these materials could *possibly* be used to make a detector for the light from a GaAs laser?
- (b) Which of the following ternary compounds could *possibly* be used to make a detector for the light from a GaAs laser?



- (c) If a reverse biased Si photodiode with a quantum efficiency of $\eta = 80\%$ and an area of 1 cm^2 is uniformly illuminated with the light from a GaAs laser to an intensity of 10 mW/cm^2 what is the photocurrent which flows?
- 17.5 (a) Determine the maximum value of the energy gap which a semiconductor, used as a photodetector, can have if it is to be sensitive to light of wavelength $\lambda_0 = 0.600 \mu\text{m}$.
- (b) A photodetector whose area is $5 \times 10^{-6} \text{ m}^2$ is irradiated with light whose wavelength is $\lambda_0 = 0.600 \mu\text{m}$ and intensity is 20 W/m^2 . Assuming each photon generates one electron-hole pair, calculate the number of pairs generated per second.
- (c) By what factor does the answer to (b) change if the intensity is reduced by a factor of $1/2$?
- (d) By what factor does the answer to (b) change if the wavelength is reduced by a factor of $1/2$?
- 17.6 We wish to design a waveguide photodiode in GaAs, with the geometry as shown in Fig. 17.5, for operation at $\lambda_0 = 0.900 \mu\text{m}$ wavelength.
- (a) If the photodiode has a waveguide thickness and depletion width $W = 3 \mu\text{m}$ at an applied reverse bias voltage of 40.5 V , what is the magnitude of the change in effective bandgap due to the electric field? (Assume $m^* = 0.067m_0$).
- (b) What length L is required to produce a quantum efficiency of 0.99 ? (Assume that scattering loss and free carrier absorption are negligible).



- 17.7 We wish to design an opto-isolator OIC consisting of a laser, waveguide, and detector monolithically coupled as shown above.
- The following assumptions apply:
- (1) All coupling of energy from input to output is via photons emitted by the laser and detected by the detector. Electrical leakage currents are negligible.
 - (2) Neglect optical coupling losses between the laser/waveguide and waveguide/detector.
 - (3) The current source driving the laser has a maximum current = 5 Amps.

- (4) It is required that the current transfer ratio be

$$\frac{I_{\text{out}}}{I_{\text{in}}} \geq 0.1.$$

Question

Which combinations(s) of lasers (a, b, c), waveguides (d, e), and detectors (f, g) described in the following paragraphs would be suitable? For each possible combination calculate $I_{\text{out}}/I_{\text{in}}$. Show all work (calculations) to indicate why certain combinations are possible or not.

Possible lasers

- (a) A DFB GaAs/(GaAl)As laser with grating spacing $\Lambda = 3400 \text{ \AA}$, index in the light emitting layer $n = 3.6$, operating in the third-order reflection mode.
- (b) An unconfined field GaAs laser with Fabry-Perot type endface reflectors. The following parameters have been either measured or established.
 - (1) Half-power points of the emission peak for spontaneous emission have been measured for this material at room temperature to be at 9200 \AA and 8800 \AA .
 - (2) Index of refraction = 3.3
 - (3) Thickness of light emitting layer = 10 microns
 - (4) Thickness of active (inverted pop.) layer = 1 micron
 - (5) Internal quantum efficiency = 0.7 (above threshold)
 - (6) Average absorption coefficient = 30 cm^{-1}
 - (7) Width = $300 \mu\text{m}$
 - (8) Length = $935 \mu\text{m}$
 - (9) Carrier energy distribution factor $\xi = 1$
 - (10) Reflectivity of Fabry-Perot surfaces = 0.4
 - (11) Negligible series $I^2 R$ loss
 - (12) Emission wavelength $\lambda_0 = 9000 \text{ \AA}$
- (c) A confined field GaAs laser with all parameters the same as in (b) except for the following:
 - 3,4) Thickness of the light emitting layer = thickness of active layer = $1 \mu\text{m}$
 - 6) Average absorption coefficient = 10 cm^{-1}

Possible Waveguides

- (d) $\text{Ga}_{0.9}\text{Al}_{0.1}\text{As}$ on $\text{Ga}_{0.8}\text{Al}_{0.2}\text{As}$ (heteroepitaxial type) thickness of the $\text{Ga}_{0.9}\text{Al}_{0.1}\text{As}$ waveguiding layer = $1.0 \mu\text{m}$, length = 5 mm
- (e) Carrier concentration reduction type in GaAs
 substrate $n = 2 \times 10^{18}/\text{cm}^3$
 waveguide $n = 1 \times 10^{15}/\text{cm}^3$
 thickness of the waveguiding layer = $3 \mu\text{m}$
 length = 2 mm

Note: Use graphs given in Chap. 4 for determining waveguide index of refraction and attenuation (absorption loss).

Possible Detectors

- (f) GaAs schottky barrier avalanche type
 - depletion width = $3 \mu\text{m}$
 - length = 0.1 mm
 - photomultiplication factor $M_{\text{ph}} = 2$
- (g) GaAs Franz-Keldysh type
 - depletion width = $3 \mu\text{m}$
 - length = 0.1 mm
 - applied electric field in depletion layer = $2 \times 10^7 \text{ V/m}$

Note: Assume all absorption in the detectors results in carrier generation, i.e. the quantum efficiency equals 1.

- 17.8 A GaInAs avalanche photodiode is used to detect the pulses of light emitted by a GaInAsP laser at a wavelength of $1.3 \mu\text{m}$. The photodiode has a quantum efficiency of 0.85 at this wavelength when operated at low reverse bias voltage, and has a photomultiplication factor of 10^3 when properly biased at avalanche breakdown. The area of the photodiode is 10^{-4} cm^2 .

What photocurrent will be produced when the photodiode is biased in the avalanche regime and a light intensity of 10^{-3} W/cm^2 from the laser is falling on it?

- 17.9 The quantum efficiency of an InGaAsP/InP avalanche photodiode is 80% when detecting $1.3 \mu\text{m}$ wavelength radiation and biased at low voltage so that no avalanche multiplication is occurring. When biased at higher voltage for avalanche detection, and incident optical power of $1.0 \mu\text{W}$ (at $\lambda = 1.3 \mu\text{m}$) produces an output photocurrent of $20 \mu\text{A}$.

- (a) What is the photomultiplication factor (avalanche gain)?
- (b) If the responsivity of the diode falls off rapidly for wavelengths greater than $1.7 \mu\text{m}$ what is the energy bandgap in the absorbing region of the diode? At $1.7 \mu\text{m}$ what is the energy bandgap in the absorbing region of the diode?

- 17.10 What key advantage does the p-i-n photodiode have over the conventional p-n junction photodiode? What key advantage does the avalanche photodiode have?

- 17.11 A silicon p-i-n photodiode is illuminated by 100 nW of light with a wavelength of $1.0 \mu\text{m}$. The quantum efficiency of the device is 55% and its dark current at the bias voltage at which it is operated is negligible. What is the magnitude of the resulting photocurrent?

References

1. D.P. Schinke, R.G. Smith, A.R. Hartmann: Photodetectors, in *Semiconductor Devices for Optical Communication*, H. Kressel, 2nd edn., Topics Appl. Phys., Vol. 39 (Springer, Berlin, Heidelberg 1982) Chap. 3

2. R.J. Keyes (ed.): *Optical and Infrared Detectors*, 2nd edn., Topics Appl. Phys., Vol. 19 (Springer, Berlin, Heidelberg 1980)
3. S.M. Sze: *Physics of Semiconductor Devices*, 2nd edn. (Wiley, New York 1981) p. 665
4. D. Bossi, R. Ade, R. Basilica, J. Berak: IEEE Photon. Tech. Lett. **5**, 166 (1993)
5. M. Erman, Ph. Riglet, Ph. Jarry, B. Martin, M. Renaud, J. Vinchant, J. Cavaillès: IEE Proc. Pt. J. Optoelectron. **138**, 101 (1991)
6. L. Giraudet, F. Banfi, S. Demiguel, G. Herve-Gruyer: Optical design of evanescently coupled, waveguide-fed photodiodes for ultrawide-band applications. IEEE Photonics Tech. Lett. **11**, 111 (1999)
7. B.G. Streetman: *Solid State Electronic Devices*, 4th edn. (Prentice-Hall, Englewood Cliffs, NJ 1994) pp. 183–190
8. S. Miller: Phys. Rev. **99**, 1234 (1955)
9. H. Melchior, W.T. Lynch: IEEE Trans. ED-**13**, 829 (1966)
10. P. Yuan, O. Baklenov, H. Nie Jr., A.L. Holmes, B.G. Streetman, J.C. Campbell: Highspeed and low-noise avalanche photodiode operating at 1.06 μm. IEEE J. Selected Topics Quant. Electr. **6**, 422 (2000)
11. I. Watanabe, T. Nakata, M. Tsuji, K. Makita, T. Torikai, K. Taguchi: High-speed, highreliability planar-structure superlattice avalanche photodiodes for 10-Gb/s optical receivers. IEEE J. Lightwave Tech. **18**, 2200 (2000)
12. K. Kato, A. Kozen, Y. Muramoto, Y. Itaya, T. Nagatsuma, M. Yaita: IEEE Photon. Tech. Lett. **6**, 719 (1994)
13. T. Hsiang, S. Alexandrou, C. Wang, M. Liu, S. Chou: SPIE Proc. **2022**, 76 (1993)
14. I.R. Mactaggart, M. Bendett, S. Tayler: IEEE J. SSC-**28**, 1018 (1993)
15. V. Hurm, M. Ludwig, J. Rosenzweig, W. Benz, M. Berroth, R. Bosch, W. Bronner, A. Hulsmann, K. Kohler, B. Raynor, J. Schneider: Electron. Lett. **29**, 9 (1993)
16. M. Leary, J. Ballantyne: IEEE Cornell Conf. on Advanced Concept High Speed Semiconductor Devices and Circuits, Ithaca, NY (1993) Proc. p. 383
17. Y. Baeyens, A. Leven, W. Bronner, V. Hurm, R. Reuter, K. Kohler, J. Rosweig, M. Schlechtweg: Millimeter-wave long-wavelength integrated optical receivers grown on GaAs. IEEE Photonics Tech. Lett. **11**, 868 (1999)
18. W. Gao, A. Khan, P. Berger, R.G. Hunsperger, G. Zydzik, H. O'Bryan, D. Sivco, A. Cho: Appl. Phys. Lett. **65**, 1930 (1994)
19. J-H. Cha, J. Kim, C-Y. Kim, S-H. Shin, Y-S. Kwon: Monolithic integration of InP-based HEMT and MSM photodiode using InGaAsP ($\lambda=1.3 \mu\text{m}$) buffer, Jpn. J. Appl. Phys. **44**, 2549 (2005)
20. D. Ostrowsky, R. Poirier, L. Reiber, C. Deverdun: Appl. Phys. Lett. **22**, 463 (1973)
21. S. Koike, H. Takahara, K. Katsura: Electron. Commun. Jpn., Pt. II: Electron. **75**, 41 (1992)
22. G. Stillman, C.M. Wolfe, I. Melngailis: Appl. Phys. Lett. **25**, 36 (1974)
23. T. Kimura, K. Daikoku: Opt. Quant. Electron. **9**, 33 (1977)
24. S. Miura, H. Kuwatsuka, T. Mikawa, O. Wada: IEEE J. LT-**6**, 399 (1988)
25. S. Chandrasakhar, J.C. Campbell, A.G. Dentai, C.H. Joyner, G.J. Qua, W.W. Snell: IEEE Electron. Dev. Lett. **8**, 512 (1987)
26. S. Chandrasekhar, J.C. Campbell, A.G. Dentai, G.J. Qua: Electron. Lett. **23**, 501 (1987)
27. S. Demiguel, N. Li, X. Li, X. Zheng, J. Kim, J.C. Campbell, H. Lu, A. Anselm: Very high-responsivity evanescently coupled photodiodes integrating a short planar multimode waveguide for high-speed applications, Photonics Technol. Lett. **15**, 1761 (2003)
28. H. Stoll, A. Yariv, R.G. Hunsperger, G. Tangonan: Appl. Phys. Lett. **23**, 664 (1973)
29. H.J. Stein: In *Int'l Conf. on Ion Implantation of Semiconductors and Other Materials*, Yorktown Heights, NY, 1972 (Plenum, New York 1973)
30. M.K. Barnoski, R.G. Hunsperger, R.G. Wilson, G. Tangonan: J. Appl. Phys. **44**, 1925 (1973)
31. S. Valete, G. Labrunie, J. Deutsch, J. Lizet: Appl. Phys. **16**, 1289 (1977)
32. D.L. Spears, A.J. Strauss, S.R. Chinn, I. Melngailis, P. Vohl: OSA/IEEE Topical Meeting on Integrated Optics, Salt Lake City, UT (1976) Digest Paper TUD3-1

33. A. Carenco, L. Menigaux: *Appl. Phys. Lett.* **37**, 880 (1980)
34. J.I. Pankove: *Optical Processes in Semiconductors* (Prentice-Hall, Englewood Cliffs, NJ 1971) p. 29
35. K.H. Nichols, W.S.C. Wang, C.M. Wolfe, G.E. Stillman: *Appl. Phys. Lett.* **31**, 631 (1977)
36. R.G. Hunsperger: Monolithic dual mode emitter-detector terminal for optical waveguide transmission lines. US Patent No. 3,952,265 (issued 20 April 1976)
37. J. Park, S. Wadekar, R.G. Hunsperger: *SPIE Proc.* **835**, 283 (1987)
38. S. Wadekar, E. Armour, M. Donhowe, R.G. Hunsperger: *SPIE Proc.* **994**, 133 (1988)
39. R. Link, K. Reichmann, T. Koch, V. Koren: *IEEE Phot. Tech. Lett.* **1**, 278 (1989)
40. S. Sakano, H. Inoue, H. Nakamura, T. Matsumura: *Electron. Lett.* **22**, 594 (1986)
41. H. Inoue, S. Tsuji: *Appl. Phys. Lett.* **51**, 1577 (1987)
42. D.S. Shin, G.L. Li, C.K. Sun, S.A. Pappert, K.K. Loi, W.S.C. Chang, P.K.L. Yu: Optoelectronic RF signal mixing using an electroabsorption waveguide as an integrated photodetector/mixer. *IEEE Photonics Tech. Lett.* **12**, 193 (2000)
43. H. Ito, T. Ohno, H. Fishimi, T. Furuta, S. Kodama, T. Ishibashi: 60 GHz high output power uni-travelling-carrier photodiodes with integrated bias circuit, *Electron. Lett.* **36**, 747 (2000)
44. D. Wolf (ed.): *Noise in Physical Systems*, Springer Ser. Electrophys., Vol. 2 (Springer, Berlin, Heidelberg 1978)
45. M. DiDomerrico Jr., O. Svelto: *IEEE Proc.* **52**, 136 (1964)
46. B. Saleh: *Photoelectron Statistics*, Springer Ser. Opt. Sci., Vol. 6 (Springer, Berlin, Heidelberg 1978)
47. J.W. Goodman, E. Rawson: Speckle phenomena in optical communication, in *Laser Speckle and Related Phenomena*, J.C. Dainty, (ed.), 2nd edn., Topics Appl. Phys., Vol. 9 (Springer, Berlin, Heidelberg 1982)

Chapter 18

Quantum-Well Devices

In all of the devices discussed in previous chapters the dimensions of device structures were large compared to the wavelength of electrons in the device. When the dimensions of the structure are reduced to the point at which they are approaching the same order of magnitude as the electron wavelength some unique properties are observed. This is the case with a class of devices that have come to be known as “quantum well” devices, which feature very thin epitaxial layers of semiconductor material. This chapter will introduce the basic concepts of quantum wells and will describe some of the novel kinds of devices that can be made by using them. Improved lasers, photodiodes, modulators and switches can all be made by employing quantum well structures. Quantum well devices can be monolithically integrated with other optical and electronic devices to produce optical integrated circuits and opto-electronic integrated circuits.

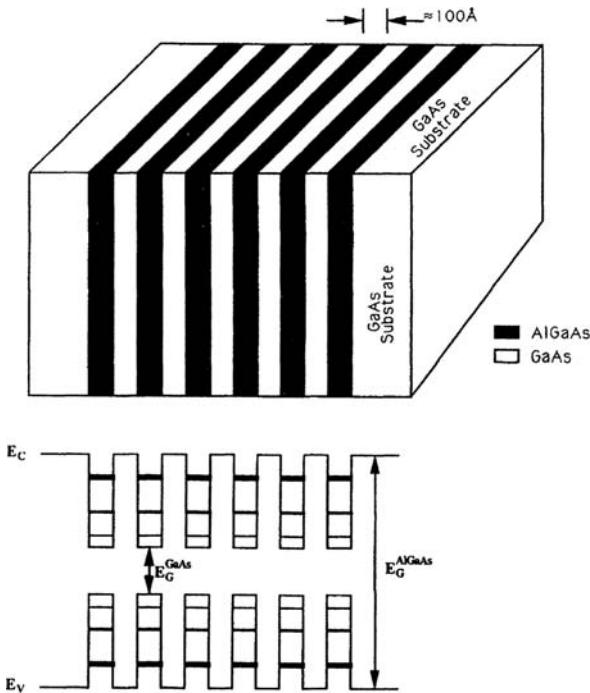
18.1 Quantum Wells and Superlattices

A quantum well structure consists of one or more very thin layers of a relatively narrow bandgap semiconductor interleaved with layers of a wider bandgap semiconductor, as shown in Fig. 18.1. The thickness of the layers is typically 100 Å, or less. An arrangement like that shown in Fig. 18.1, with many layers, is called a “Multiple Quantum Well” (MQW) structure. Single Quantum Well (SQW) structures, with one narrow bandgap layer, are also useful. The required thin layers can be grown by either MBE or MOCVD techniques such as those described in Chapter 4. A typical MQW structure might have about 100 layers. The GaAs-AlAs material system is particularly convenient for the growth of MQW devices because GaAs and AlAs have almost identical lattice constants and thus interfacial strain can be avoided. However, GaInAsP is also a suitable material, as long as the concentrations of the constituent elements are properly chosen to produce lattice matching. (In the case of high power lasers, it is sometimes beneficial to purposely produce a strained-layer structure to reduce losses, as is explained in Section 18.2.)

The special properties of these very thin layers result from the confinement of carriers (electrons and holes) in a manner analogous to the well-known quantum

Fig. 18.1

Multiple-quantum-well (MQW) structure



mechanical problem of the “particle in a box” [1]. In this case, the carrier is confined to the narrow bandgap “well” by the larger bandgap “barrier” layers. The magnitude of the wavefunction of the electron (or hole) must approach zero at the barrier wall because the probability of finding the particle within the wall is very small. (If the wall were infinitely high, the wave amplitude would have to equal zero at the barrier.) Hence, the wavefunction must form a standing wave pattern, sinusoidally varying within the well and damping to near zero at the edge of the barrier. The set of wavefunctions which satisfy these boundary conditions corresponds to only certain allowed states for the carrier. The carrier motion is thus quantized, with discrete allowed energies corresponding to the different wavefunctions.

In the MQW structure described in the previous paragraphs it was assumed that the barrier layers were sufficiently thick that the tail of the wavefunction penetrated only slightly into the barrier layer. Carriers are thus confined to a particular layer and act independently of carriers in the other layers. If a structure is fabricated with many wells and barrier layers, but the barrier layers are made so thin that the electron or hole wavefunction can penetrate the barrier, there is coupling between the allowed states. In that case, the electrons and holes are not clearly confined to a particular state associated with a given well. Instead they exist in a coupled system of states with the wavefunction extending over many wells. Thus the behavior of the carriers is influenced by the long-range periodic modulation superimposed upon

the crystalline potential of the host material. The similarity between this situation and that which normally exists for a charge carrier in a crystalline lattice has led to the use of the term “superlattice” to describe such MQW structures with very thin barriers. The formation of a superlattice actually results in the creation of a new material, with a bandgap and electronic and optical properties which may be significantly different from those of the host material.

Both superlattices and quantum well structures can be fabricated in a variety of semiconductor materials. GaAs wells with GaAlAs barriers are generally used for wavelengths near $0.9 \mu\text{m}$, while InGaAsP can be used for 1.3 and $1.5 \mu\text{m}$ wavelengths. Many other III-V and II-VI materials are also suitable. The use of quantum wells and superlattices in opto-electronic devices is a relatively new field of endeavor, and much research is presently being conducted. In the following sections of this chapter some of the more notable examples of quantum well devices are described, along with the relevant theory of operation.

18.2 Quantum-Well Lasers

In a quantum well laser the quantization of the allowed electron and hole energies into discrete levels within the energy well, as shown in Fig. 18.1, reduces the total number of carriers needed to achieve a given level of population inversion. The free carrier absorption coefficient, which is proportional to the number of carriers, is also reduced. As a consequence, the threshold current density is reduced by approximately a factor of 10 as compared to that of a conventional double heterostructure laser diode.

18.2.1 Single-Quantum-Well Lasers

Only a single-quantum well is required in order to realize the benefits of quantization of the electron and hole states. If it is assumed that the potential well is deep (i.e., that the barrier height is infinite) the solutions of Schrödinger's equation in the n th quantized level is given by [2]

$$E_n = \frac{\hbar^2}{2m_e} \left(\frac{n\pi}{L_z} \right)^2 + \frac{\hbar^2}{2m_e} (k_x^2 + k_y^2), \quad n = 1, 2, 3, \dots, \quad (18.1)$$

where the z direction is taken normal to the thin layer which forms the well, m_e is the effective mass, L_z is the thickness of the layer and k_x , k_y and k_z are the magnitudes of the wavevector. A corresponding relation for the energy of a hole in the n th level can be obtained by merely replacing m_e with m_h , the effective mass of a hole, and remembering that increasing energy (hence n) for a hole is in the downward direction in an electron energy diagram such as that shown in Fig. 18.1. Since

$$k_i = \frac{n\pi}{L_i}, i = x, y, z \quad \text{and} \quad L_z \ll L_x, L_y, \quad (18.2)$$

the energy levels in the x and y directions will be closely spaced, and for each allowed value of E_n a two-dimensional energy band will exist in the x - y plane with a density of states that is independent of energy. Thus, for increasing energy, the cumulative density of states function will exhibit a sharp step at each allowed value of E_n . The magnitude of this step $\Delta\rho(E)$ is given by [3]

$$\Delta\rho(E) = \frac{m}{\pi\hbar^2}. \quad (18.3)$$

where m is the effective mass of the electron or hole, as the case may be.

The assumption of an infinite barrier height in the derivation of (18.1) and (18.3) is not unduly restrictive. For most device applications, while the wells are finite depth, they are sufficiently deep that only a small perturbation of the wavefunction results as compared to that of the infinite-barrier case. The conduction and valence bands are still quantized into two-dimensional subbands with step-like densities of states. Since photon generation in a quantum well laser occurs through electron transitions from states in one of these subbands of the conduction band to states in one of the subbands of the valence band, the gain curve as a function of photon energy, ω , has steps which occur at photon energies

$$\hbar\omega = E_g + E_{nc} + E_{nv}, \quad (18.4)$$

where E_g is the bandgap energy and E_{nc} and E_{nv} are the energies of the n th levels in the conduction and valence bands, respectively. Plotted curves of gain vs. photon energy calculated by M. Mittelstein have been published by Yariv [4]. The estimated maximum gain of a single quantum well laser operating in the $n = 1$ transition is $g \simeq 100 \text{ cm}^{-1}$. However, in a practical laser, that gain must be multiplied by a confinement factor to account for the fact that the laser mode does not travel entirely within the active region. The confinement factor Γ_a is given by [5]

$$\Gamma_a = \frac{\int_{-L_z/2}^{L_z/2} |E|^2 dz}{\int_{-\infty}^{\infty} |E|^2 dz} \simeq \frac{L_z}{W}, \quad (18.5)$$

where W is the width of the mode. Despite some loss of the potentially available gain because of modal spreading, the reduction in the number of carriers needed for population inversion, and reduced free carrier absorption, result in about a factor of 10 decrease in the threshold current density of the quantum well laser as compared to that of a conventional laser. The quantization of electron and hole allowed energy states also results in a decrease in the linewidth of the emitted radiation.

Quantum wells can be combined with many of the other structures that have been developed over the years to improve laser performance. For example, Fig. 18.2 shows the cross-section of a ridge-waveguide, graded index, separate confinement

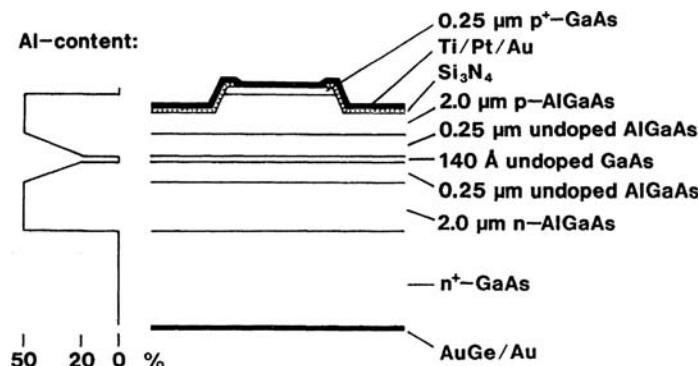


Fig. 18.2 GaAs/GaAlAs single quantum well laser [6]

heterostructure, single quantum well laser [6]. This GRIN-SCH-SQW structure yielded a laser with a threshold current density of 20 mA, and a linewidth of only 1.5 MHz at an output power of 5 mW. The sidemode suppression was measured to be greater than 24 dB.

The structure shown in Fig. 18.2 was grown by the MOCVD method, at atmospheric pressure. The SQW was 140 \AA thick, and was surrounded by linearly graded GaAlAs layers for efficient carrier collection [7] and improved optical confinement [8]. The optical confinement factor for this structure was calculated to be 4% [9]. Ridge waveguides 5 μm wide and 2 μm deep were fabricated by conventional wet chemical etching, and lasers with cavity lengths ranging from 300 to 800 μm were cleaved and mounted p-side down on heatsinks.

In operation the lasers exhibited a differential quantum efficiency of 66% at a temperature of 20°C, controlled to within 0.01°C by a Peltier cooler. The emitted wavelength was 872 nm, and a single longitudinal mode was observed for output power as high as 5 mW. As mentioned previously, the linewidth of the emitted light was unusually narrow. The linewidth-power product for a laser with cavity length $L = 800 \mu\text{m}$ was determined to be 6.4 MHz mW, with the narrowest linewidth of 1.5 MHz being measured at an output power of 5 mW. The linewidth power product was found to follow an L^2 dependence on cavity length, L . This laser demonstrates the advantages of combining the GRIN-SCH structure, for enhanced confinement characteristics, and the ridge waveguide, for low threshold current and single mode oscillation, with a long-cavity-length, SQW structure for further reduction of threshold current and linewidth of emission.

Single quantum well lasers with threshold currents less than 1 mA have been produced by adding reflective coatings to the front and rear facets of a GRIN-SCH-SQW structure [10]. A minimum threshold current of 0.95 mA was achieved with facet reflectivities of 70%, a cavity length of 250 μm , and an active region stripe width of 1 μm . The SQW in that case was a 100 \AA thick layer of GaAs surrounded by 0.2 μm thick, undoped, graded layers of $\text{Ga}_{0.8}\text{Al}_{0.2}\text{As} \rightarrow \text{Ga}_{0.5}\text{Al}_{0.5}\text{As}$. An identical laser, but with uncoated facets, had a threshold current of 3.3 mA. When

dielectric coatings ($1/4\lambda\text{Si}$, $1/4\lambda\text{Al}_2\text{O}_3$) were subsequently added to the facets of this control laser, its threshold current was reduced to 0.95 mA. The addition of reflective coatings to the facets of an SQW laser has a much greater effect on its threshold current than the same process would have on that of a conventional DH laser. This is due to the fundamentally lower current density needed to render the thin active layer transparent.

Low threshold current SQW lasers have also been made in InGaAs-AlGaAs/GaAs by using MOCVD and selective area growth to define a 3 μm stripe width, 446 μm long, BH laser [11]. These devices had a threshold current of 2 mA and a differential quantum efficiency of 91%.

While the SQW laser diode is a relatively new device, its fabrication has not been confined to just the laboratory, SQW laser became available as commercial products as early as 1987.

18.2.2 Multiple Quantum Well Lasers

Since the gain of a single quantum well laser operating in the $n = 1$ transition is limited to about 100 cm^{-1} , it is often advantageous to use many well and barrier layers to form a multiple quantum well laser. At sufficiently high injection current levels so that the quasi Fermi levels cross over the $n = 1$ levels in both the conduction and valence bands, the gain of an N well structure is just N times that of a single quantum well. However, at lower drive current levels, the gain of the MQW structure is actually less than that of the single quantum well.

While a higher level of drive current density is required to optimize the gain of an MQW laser, that doesn't mean that a low threshold current cannot be achieved. For example, Tsang [12] has reported MQW lasers with a threshold current of only 2.5 mA. In that case the laser featured a 4 well structure with 208 Å thick GaAs wells and 76 Å thick barriers of $\text{Ga}_{(1-x)}\text{Al}_x\text{As}$. The optimum aluminum concentration in the barriers was shown to be $x \simeq 0.2$. Greater concentration of Al in the barrier layers was found to unduly inhibit the injection of carriers because of the greater barrier height.

MQW lasers exhibit less temperature dependence of the threshold current than do conventional DH lasers. The temperature dependence of threshold current density can be fitted to the relation [13]

$$J_{\text{th}}(T) = J_{\text{th}}(0) \exp\left(\frac{T}{T_0}\right), \quad (18.6)$$

where the constant T_0 gives a measure of the degree of temperature insensitivity. For conventional DH lasers T_0 is usually in the range from 120 to 165 °C, while for MQW lasers T_0 may be over 400 °C. Of course, the value of T_0 depends strongly on the particular MQW structure used and on the specific geometry of the laser diode. However, values of T_0 reported for MQW lasers generally are significantly greater than those for conventional DH lasers. For example, Chin et al. [13] have measured

T_0 for stripe geometry quantum well lasers grown by MOCVD and found $168^\circ\text{C} \leq T_0 \leq 437^\circ\text{C}$. Diodes which emitted at longer wavelengths were found to exhibit larger values of T_0 (in the wavelength range 8435–8650 Å).

The emission wavelength of a quantum well laser depends strongly on the dimensions of the quantum well structure, since electron transitions occur between the subbands of the conduction band and those of the valence band. Xu et al. [14] have measured the emission wavelengths of MOCVD-grown MQW lasers in the temperature range from 77 K to 300 K and found a close correspondence to the subband energy levels and to the bandgap energy. In the emission spectra shown in Fig. 18.3 the dominant emission peak corresponds to the $n = 1$ subband and the smaller peak corresponds to the $n = 2$ subband state. The fact that the emission wavelength also depends on the dimensions of the quantum well structure is shown in Fig. 18.4a, in which the emission spectra for two InGaAsP MQW lasers are displayed [15]. The lasers are essentially the same except that sample #1 has 3 well layers with $L_z = 200\text{ \AA}$ and sample #2 has 3 well layers with $L_z = 100\text{ \AA}$. A significant difference in the emitted wavelengths is noted due to the quantum-size shift effect [16].

The lasers for which the spectra of Fig. 18.4a were measured also serve as a good illustration of the narrow linewidth obtainable from MQW lasers. These are BH stripe-geometry lasers with a stripe width of about $2.5\text{ }\mu\text{m}$ and a cavity length of $300\text{ }\mu\text{m}$. Notably, they were fabricated by liquid phase epitaxy (LPE) rather than by

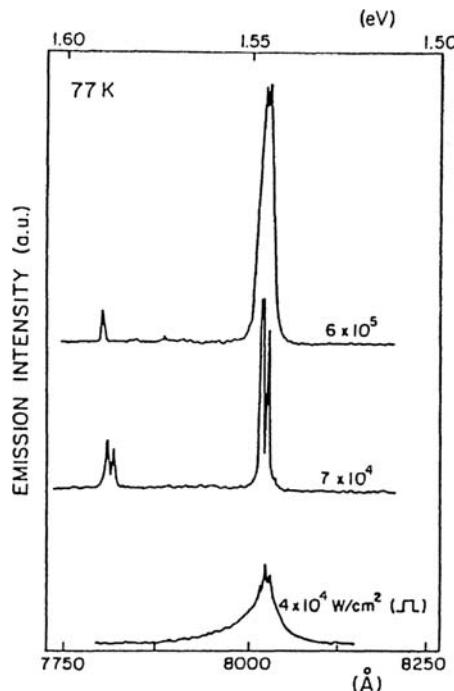
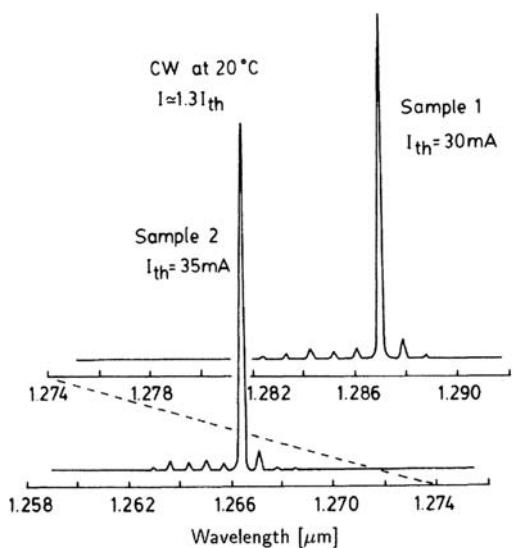
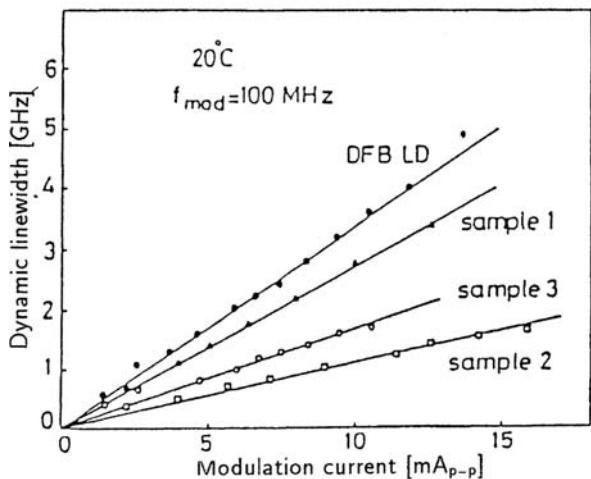


Fig. 18.3 Emission spectra of an MQW laser [4]

Fig. 18.4 a Emission spectra for two InGaAsP MQW lasers [15].



a



b

MOCVD or MBE. A rapid sliding growth technique was used at a low growth temperature (589°C) [17]. The dynamic spectral linewidths of these lasers were measured for a modulation frequency of 100 MHz. The dynamic linewidth is particularly important in regard to long distance telecommunications because of the problem of frequency shift during pulsing, or “chirping”, which was described in Chapter 16. The modulation current dependence of dynamic spectral linewidth is shown in Fig. 18.4b. Samples 1 and 2 are as in Fig. 18.4a; sample #3 is identical except it has

$L_z = 150 \text{ \AA}$. The dynamic linewidth of a conventional DFB laser is also shown in the figure to permit easy comparison. It can be seen that the dynamic linewidth of the MQW lasers decreases as the quantum well width decreases and that it is less than that of the DFB laser for all cases shown. The lasers were DC biased at $1.3l_{\text{th}}$ for these measurements. The reduced chirping in the MQW lasers can be attributed to reduction of the linewidth enhancement factor α and the confinement factor Γ_a [18]. The minimum spectral linewidth of an unmodulated DFB-MQW laser is inherently very narrow, as demonstrated by Okai et al. [19] who have measured a spectral linewidth of 56 KHz for one at an output power of 25 mW. White noise was found to be the limiting factor.

In the growth of DH lasers and MQW lasers the usual practice is to control the compositions of adjacent layers appropriately so as to produce lattice matching and eliminate interfacial strain. However, recent work has shown that in the case of high power lasers, it may be desirable to purposely produce a controlled amount of strain. Strained InGaAs layers grown on InP or GaAs have a modified valence band structure [20] and an increased conduction band discontinuity [21], which it is believed may eliminate Auger recombination and intervalence-band absorption that otherwise cause optical losses [22]. A CW output power of 120 mW has been reported for $1.5 \mu\text{m}$ wavelength modulation-doped InGaAs strained-layer quantum well lasers [22]. In these devices, four 30 \AA thick $\text{In}_{0.8}\text{Ga}_{0.2}\text{As}$ wells with 1.8% strain separated by three 200 \AA thick InGaAsP barrier layers were grown on an InP substrate by low pressure MOCVD. The cavity length was $750 \mu\text{m}$. A separate confinement, buried heterostructure was combined with the strained MQW structure. Kink free CW operation was observed at output power levels as high as 120 mW per facet. For pulsed operation with a 400 ns pulse width at a 40 kHz repetition rate, an output power of 360 mW per facet was obtained. From measurements of the differential efficiency of devices with different lengths, a cavity loss of $\alpha = 13 \text{ cm}^{-1}$ and an internal quantum efficiency of $\eta_i \simeq 100\%$ were determined, indicating reduced optical losses.

Strained-layer quantum well heterostructures can also be fabricated in the InGaAs-GaAs-AlGaAs system to produce lasers emitting in the 0.9 to $1.3 \mu\text{m}$ wavelength range. High power diode lasers operating in that range are attractive as small, efficient, low-cost optical pumping sources for rare earth doped glass lasers. Coleman [23] has reported stripe geometry, buried heterostructure lasers of this type that exhibit very low threshold currents (7 – 9 mA), stable fundamental optical mode properties to greater than 30 times the threshold current and high power operation ($> 130 \text{ mW}/\text{uncoated facet}$). By combining devices of this type in a monolithic parallel array having a cavity length equal to $440 \mu\text{m}$ and a width of 1.6 mm , an output power of 2.5 W per uncoated facet was obtained at a wavelength of $1.03 \mu\text{m}$. The far field emission pattern of the array indicated that the lasers were phase locked.

Low-threshold-current and high-temperature operation of $1.3 \mu\text{m}$ wavelength AlGaInAs/AlGaInAs strain-compensated multiple-quantum-well laser diodes has been achieved [24]. These linearly-graded-index, separate-confinement-heterostructure lasers have threshold current density and differential quantum efficiency of 400 A/cm^2 and 22% , respectively, for as-cleaved broad-area lasers with a

900 μm cavity length. When etched to produce 3- μm -ridge-stripe laser diodes without facet coating the threshold current and slope efficiency are 12 mA and 0.17 W/A, respectively.

The concept of the quantum well laser, in which carriers are narrowly confined in one dimension, can be extended to the two and three-dimensional analogues of the “quantum wire” and “quantum dot” lasers [25]. For example, in a quantum dot laser, the electron would be confined in all three dimensions to a distance on the order of its wavelength in the crystal. One could picture GaAs dots with a GaAlAs host. Theoretical analyses predict that quantum dot or quantum wire lasers would have lower threshold currents, lower noise performance and higher modulation speeds than those of even quantum well lasers [26]. The fabrication of quantum wire and quantum dot lasers has been made possible by the development of new fabrication techniques in the field of what has come to be called nanotechnology. (See Chapter 22.) One approach is to begin with a substrate containing a quantum well structure, and then to define lateral dimensions by electron or ion beam lithography coupled with either etching [27] or ion implantation [28]. A more sophisticated approach is to produce localized compositional control through MBE growth on GaAs substrates [29].

The advent of quantum well structures has changed the direction of semiconductor laser research and development over a period of just a few years. Substantial improvements in threshold current density, linewidth and temperature sensitivity have already been demonstrated, even as compared to the properties of the DFB laser [25]. Combining DFB and MQW structures has produced devices in InGaAsP with linewidth-power products of 1.9–4.0 with minimum linewidths of 1.8–2.2 MHz [30]. Some quantum well lasers already have found their way to the commercial market. It seems certain that continued development of this relatively new device, including the development of quantum wire and quantum dot lasers, will lead to further improvements in the operating characteristics of semiconductor laser diodes.

18.3 Quantum-Well Modulators and Switches

The same type of quantum well structures that have proven so useful in laser diodes have also been found to improve the performance of modulators and switches. In these devices the improvement generally results from enhancement of the electro-optic and electro-absorption effects.

18.3.1 *Electro-Absorption Modulators*

In “bulk” semiconductor crystalline materials the absorption spectrum has a characteristic “absorption edge” at approximately the wavelength corresponding to the bandgap energy. For longer wavelengths interband absorption is absent, but the absorption coefficient α increases rapidly as wavelength is decreased below the

bandgap wavelength. For quantum well structures, because of the quantization of both electron and hole energies, the absorption spectrum has a series of steps corresponding to the $n = 1, 2, 3 \dots$ energy levels. These steps are typically 50–200 meV in width, and the first step occurs at an energy somewhat above the bandgap energy. In other words, the absorption edge is shifted to a higher energy corresponding to the $n = 1$ transition. In addition, “exciton” peaks are observed at each step. These peaks correspond to absorption of photons which create pairs of electrons and holes that are not completely separated but are orbiting around each other. The pair resembles a hydrogen atom with the hole acting as the proton. Normally, excitons have such a short lifetime that their effects can be observed only at cryogenic temperatures. However, in a quantum well, the hole and electron of the exciton pair are compressed to a closer spacing in one dimension, typically on the order of 100 Å. The spherical shape of the exciton changes to that of a three-dimensional ellipsoid. This greatly increases the coulombic bonding force and makes the exciton more stable, even at room temperature. The presence of excitons is significant because it strengthens the response and sharpens the absorption curve near the bandgap wavelength at which electro-absorption devices work.

In a bulk semiconductor material, application of an electric field induces a small shift in the exciton absorption peak to longer wavelengths because polarization reduces the binding energy. The effect is called the Stark effect by analogy with the Stark shift in the absorption lines of a hydrogen atom in the presence of an electric field. However, in a bulk semiconductor, the small Stark shift is masked by a broadening of the exciton peak due to exciton ionization, so it is not very useful.

In an electric field is applied to a quantum well structure in the direction normal to the planes of the wells, the electrons and holes of the exciton pairs are confined. The hole and electron tend to move to opposite sides of the well and the polarization reduces the binding energy of the pair, but the walls prevent the electron and hole from going too far. Thus, field ionization of the exciton is inhibited, and the sharp exciton absorption peaks are preserved. This “Quantum” Confined Stark Effect” QCSE [31] is very useful in modulators and switches because even moderate fields ($\sim 10^4$ V/cm) can cause a significant change in absorption [32]. The electro-absorption effect is approximately 50 times larger in multiple quantum well structures than it is in bulk semiconductors.

Early QCSE modulators were made in MQW layers of GaAs and GaAlAs in which the light was introduced in the direction normal to the planes of the quantum wells [31, 33, 34]. Even though an absorption change $\Delta\alpha$ as much as fifty times larger than that of bulk electro-absorption modulation was observed for drive voltages less than 10 volts, the all-important on-off ratio was limited. The on-off ratio of an electro-absorption intensity modulator is given by [35]

$$R = \exp \Delta\alpha L, \quad (18.7)$$

where $\Delta\alpha$ is the maximum achievable change in the absorption coefficient and L is the interaction length between the light and the absorptive material. In a transverse QCSE modulator, with light introduced in the direction normal to the planes of the

wells, the interaction length L is limited to the thickness of the MQW structure, typically $1 \mu\text{m}$ at most. Thus, even though $\Delta\alpha$ is large, the on-off ratio will be relatively low. Simply growing more wells in the MQW structure to increase L is not the best approach since it would increase the drive voltage required to achieve a given magnitude of electric field.

Studies of the limits to normal incidence electro-absorption modulation in GaAs/GaAlAs MQW diodes [36] have indicated that there is an optimum number of wells for a given bias voltage, which increases roughly linearly with voltage. Performance also depends strongly on the well thickness, and to a lesser extent on background doping. For example, in samples with low background doping ($\sim 10^{14}/\text{cm}^3$), optimum performance was $R = 2.1$ for 63 100-Å wells at 5 V operating voltage, and $R = 10$ for 200 100-Å wells at > 14 V operating voltage. With a well width of 50 Å, improved performance compared to that of 100-Å well devices, could be obtained, but only at operating voltages greater than 28 V. Generally an operating voltage greater than 10 V would be considered excessive.

A more effective way to increase L is to use a waveguide modulator structure such as that shown in Fig. 18.5 [37]. In this case the direction of propagation of the light is in the plane of the MQW layers. The light is confined to propagate in this plane by a waveguide structure. Two quantum wells, 94 Å thick, were placed in the center of a p-i-n diode, and light was confined in the plane of the layers by leaky superlattice (SL) waveguides. In this structure L can be made as long as necessary to produce the desired on-off ratio. Also, it is possible to introduce light with either a TE or TM polarization, which permits the fabrication of polarization selective modulators, since absorption is different for the two polarizations. In calculating the on-off ratio for a waveguide structure like that of Fig. 18.5, it is necessary to take

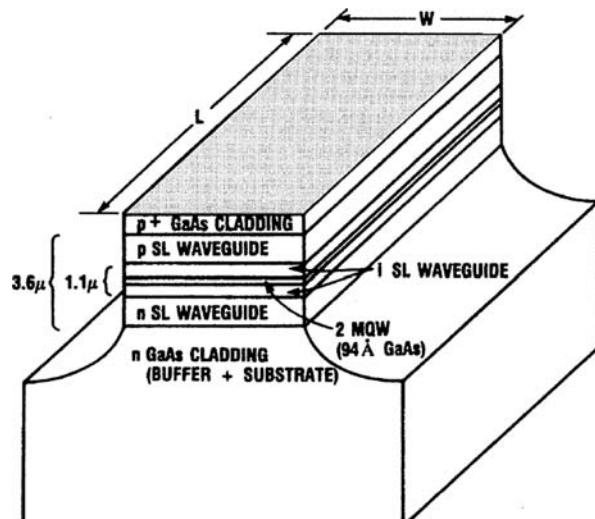


Fig. 18.5 Waveguide MQW electro-absorption modulator [37]

into account the fact that the quantum wells make up only a part of the thickness of the waveguide. Thus (18.7) becomes [38].

$$R = \exp \Gamma \Delta_\alpha L, \quad (18.8)$$

where Γ is the overlap integral between the quantum wells and the optical mode of the slab waveguide, $E_{\text{opt}}(Z)$:

$$\Gamma \equiv \frac{\int_{\text{MQW}_S} |E_{\text{opt}}|^2 dZ}{\int_{-\infty}^{\infty} |E_{\text{opt}}|^2 dZ}. \quad (18.9)$$

For the device of Fig. 18.5 the calculated value of Γ is 0.010 [39]. For a length of $L = 150 \mu\text{m}$, and a width of $W = 40 \mu\text{m}$, the measured on-off ratio was $R = 10$ for a drive voltage of 10 volts. An insertion loss of 7 dB was measured, which was attributed to endface reflections (3 dB), residual absorption in the “on” state (2 dB), and waveguide loss in the leaky waveguide structure (2 dB). Reflection loss, of course, could be greatly reduced by coating the facets with antireflecting dielectric films.

Since external modulators are generally only used at frequencies above 1 GHz, for which direct modulation of the semiconductor laser becomes difficult, the frequency response of electro-absorption modulators is an important parameter. Time domain measurements of the pulse response of the modulator shown in Fig. 18.5 have indicated a full-width-half-maximum (FWHM) pulse resolution of less than 100 ps [37]. The frequency response of MQW electro-absorption modulators is generally limited by the RC time constant of the modulator and its driver circuit. The series resistance of the MQW diode and its capacitance can both be made very small (resistance of a few ohms and capacitance of a few tenths of a picofarad). Thus operating frequencies greater than tens of GHz can be projected for the device itself. The limiting factors in most cases are the parasitic resistance, inductance and capacitance of the driver circuit, which have limited maximum frequencies. One solution to this problem lies in monolithically integrating the MQW modulator with the laser diode and with the driver circuit on a single semiconductor chip. Monolithically integrated DFB lasers and MQW electroabsorption modulators in GaInAsP have achieved modulation rates as high as 40 Gbit/s [40–42].

MQW modulators can also be used to produce high-frequency modulation in hybrid optical integrated circuits. In that case care must be taken to efficiently couple the lightwaves into and out of the QCSE modulator diode, and the modulating microwave signal must be impedance matched to the diode. An example of this type of modulator is the one described by Kuri et al. [43], which was optimized for operation at 60 GHz. The hybrid OIC is shown in Fig. 18.6.

The optical input is from the right, through a polarization-maintaining fiber. The out is on the left, into a standard single-mode fiber. The electroabsorption modulator diode contains a tensile-strained MQW of InGaAsP/InGaAsP. A voltage of 3 V

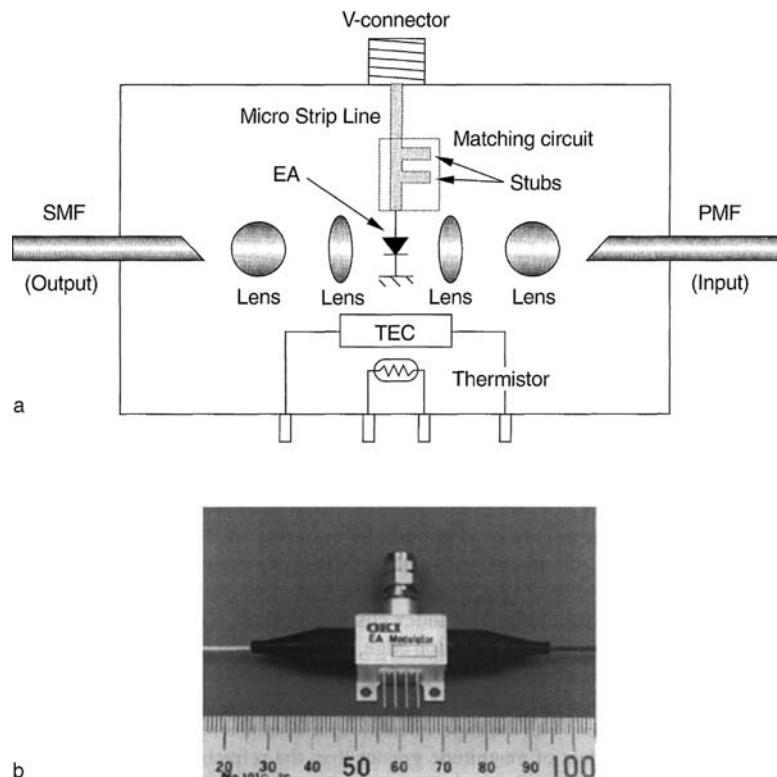


Fig. 18.6 a Configuration of modulator OIC **b** Packaged modulator module [43] © 1999 IEEE

was sufficient to produce 98% extinction. The modulating microwave signal was input via a microstrip line with a stub impedance matching circuit. A thermistor and thermoelectric cooler (TEC) were used to maintain the OIC at 25 °C. This modulator was designed for use in a fiber-optic, millimeter-wave, 60-GHz downlink system that is described in Chapter 21.

The work on longer wavelength devices will surely continue, since the major application for QCSE modulators is in long distance telecommunication systems, in which external modulators can be used to avoid high frequency chirping of directly driven laser diodes.

18.3.2 Electro-Optic Effect in Quantum Wells

The strong electro-absorption effect in quantum well devices implies that an electro-optic effect must also be present, because a change in absorption is linked with a change in index of refraction by the Kramers-Kronig relations [44]. These relations involved Cauchy integrals which permit a calculation of the index of refraction $n(\lambda)$

when the absorption spectrum $\alpha(\lambda)$ is completely known, or a calculation of the absorption coefficient $\alpha(\lambda)$ when the index spectrum $\alpha(\lambda)$ is completely known. In practical applications $n(\lambda)$ or $\alpha(\lambda)$ are usually measured over only a limited wavelength range and projections are made of their values for wavelengths extending from zero to infinity. Hiroshima [45] has used the Kramers-Kronig relations to calculate electric field induced refractive index changes in GaAs/GaAlAs quantum well structures, near the excitonic absorption edge. The calculated maximum variation in the refractive index was approximately -0.065 ($\Delta n = 1.8\%$) for an electric field of 80 kV/cm for both $60\text{-}\text{\AA}$ thick and $100\text{-}\text{\AA}$ thick GaAs/Ga_{0.6}Al_{0.4}As quantum wells. This value is consistent with the experimental value obtained by Nagai et al. [46], and is about two orders of magnitude greater than the refractive index variation in conventional III-V semiconductor bulk crystals.

The presence of a significant electro-optic effect could be a problem in MQW electro-absorption intensity modulators if it were strong enough to cause excessive phase modulation of the output light signal, resulting in chirping of the emission frequency. The relative strengths of the electro-absorption and electro-optic effects in MQW devices can be conveniently compared in terms of the “linewidth enhancement factor”, given by

$$\text{LEF} \equiv \frac{\Delta n_{\text{real}}}{\Delta n_{\text{imag}}}, \quad (18.10)$$

where Δn_{real} and Δn_{imag} are the electric field induced changes in the real and imaginary parts of the index of refraction. Thus, in a purely electro-absorptive material, $\text{LEF} = 0$, while in a purely electro-optic material $\text{LEF} = \infty$. The value of the linewidth enhancement factor has been measured by Wood et al. [47] for an MQW electro-absorption modulator, and found to be $\text{LEF} = 1$. This measurement was made at the conditions of a wavelength, bias and polarization which were optimum for intensity modulation, and hence is representative of typical modulator operating conditions. Since directly modulated lasers have values of LEF in the range of 3–6 [48], it appears that chirping of electro-absorption modulators due to the electro-optic effect will not be a serious problem. Semi-empirical calculations of the refractive index changes induced by a perpendicular electric field in the quantum wells of a QCSE modulator also have indicated that electro-absorption modulators are capable of very-low-chirp modulation [49].

While the electro-optic effect is not large enough to significantly disturb the operation of MQW electro-absorption modulators, it may be possible to use it to make MQW electro-optic phase modulators. Changes of the real part of the index of refraction in the presence of an electric field of as much as 3.7% have been reported for MQW structures [47], as compared to typical changes of only a few hundredths of a percent in bulk electro-optic materials. This large index change in MQW devices suggests that the length of electro-optic phase modulators might be significantly shortened by the use of MQWs. Of course, the modulator would have to be designed in such a way as to avoid the possible detrimental effect of electro-absorption intensity modulation.

18.3.3 Multiple Quantum Well Switches

Multiple quantum well structures can also be used advantageously in the fabrication of optical switches. The familiar dual-channel directional coupler switch described in Chapter 9 (Fig. 9.5) can be modified by using an MQW waveguide structure. Wa et al. [18.50] have made such a switch by depositing 20- μm wide gold stripe electrodes on to a waveguide consisting of a 1- μm thick MQW layer on top of a $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ layers, on top of a $\text{Ga}_{0.4}\text{Al}_{0.6}\text{As}$ layer, on a GaAs substrate. The MQW layer was composed of 25 GaAs wells, 100- \AA thick, separated by 300- \AA thick $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ barrier layers. The ratio of the thickness of the GaAs and GaAlAs layers was chosen so that the MQW layer and the $\text{Ga}_{0.7}\text{Al}_{0.3}\text{As}$ cladding layer could support only the lowest order TE and TM slab modes. The $\text{Ga}_{0.4}\text{Al}_{0.6}\text{As}$ layer was included to prevent any power from the evanescent tail of the guided mode from leaking into the heavily doped ($\simeq 10^{18}/\text{cm}^3$) substrate. By evaporating two 0.5- μm thick gold stripes on the MQW layer surface, a strain pattern was produced in the semiconductor which resulted in a spatial variation in refractive index via the photocalastic effect, and four channel waveguides were formed underneath the edges of the metal stripes. The two waveguides in the channel regions between the gold stripes (spaced 10 μm apart) formed a dual-channel directional coupler.

This optical switch was tested using an 8500 \AA GaAlAs laser diode as the light source. The critical coupling length of the coupler was found to be 1.7 mm, which is shorter than that which would be expected for similarly spaced channel waveguides without an MQW layer. Another interesting feature of this device was that it was possible to produce a phase shift of π radians, and resultant switching, through the nonlinear electro-optic effect, by increasing the optical power output of the laser. The differential phase shift $\Delta\phi$ for an input power P_i is given by [50]

$$\Delta\phi = \frac{2\pi n_2 p_i}{\lambda_0 A} \left(\frac{1 - \exp(\alpha l)}{a} \right), \quad (18.11)$$

where A is the cross-sectional area of the guided wave, α is the attenuation constant, n_2 is a nonlinear refractive index coefficient, l is the waveguide length and λ_0 is the vacuum wavelength. For the device which has been described, the measured values of the various parameters were: $A = 10 \mu\text{m}^2$, $\alpha = 15 \text{ cm}^{-1}$, $l = 2 \text{ mm}$, $\lambda_0 = 0.85 \mu\text{m}$, leading to an estimated $n_2 \sim 10^{-7} \text{ cm}^2/\text{W}$ based on observed phase shift. Optically controlled phase shift and switching such as was observed in this device are intriguing because they suggest the possibility of all-optical logic circuitry, capable of functioning without any electronic inputs or control signals. However, the required optical power level will have to be greatly reduced from that which was used in this device ($\sim 1 \text{ mW}$) before an optical computer of this type would be feasible. The electro-absorption effect, which was described previously in regard to intensity modulators, can also be used to make optical switches. A 2×2 optical gate matrix switch, monolithically integrating MQW electro-absorption switches and waveguide circuits is shown in Fig. 18.7 [51]. Arrays of such switches could be used for constructing large scale optical switching systems. In the switch shown in

Fig. 18.7 2×2 MQW optical gate matrix switch [51]

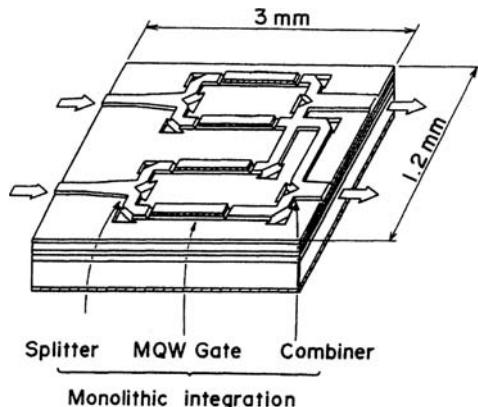


Fig. 18.7, a small chip size has been obtained by monolithically integrating MQW electro-absorption switches with miniaturized optical splitters and combiners having corner mirrors fabricated by the reactive-ion-beam-etching (RIBE) method. Electrical isolation between the four MQW gates was achieved by forming grooves around the individual gates. The MQW structure consisted of 28 GaAs quantum well layers each 80 Å thick, separated by 80-Å thick GaAlAs barrier layers, incorporated into a p-i-n diode structure.

In operation near the exciton absorption peak wavelength (approximately 838 nm) the MQW gates had an extinction ratio (crosstalk) of 20 dB for an applied voltage of 12 V (for the TE mode). The gate switching response time was measured to be 28 ps. However, the insertion loss of the switch was relatively large – a total of 24 dB, including 9 dB propagation loss, 8 dB mirror loss, 4 dB scattering loss and 3 dB splitting loss.

The MQW electro-absorption switch is attractive because it can provide a high extinction ratio at a lower applied voltage than that required by a conventional DH electro-absorption switch, because of the QCSE. A high switching speed can be achieved because of the smaller device size (and capacitance). Also, MQWs have relatively low absorption saturation. Wood et al. have shown that on-off ratios of greater than 10 dB can still be obtained at intensities as high as 500 kW/cm², which corresponds to 10.5 mW coupled into a device of 1 μm × 3 μm cross-sectional area [52]. However, the optical losses due to absorption and scattering are relatively high.

A MQW structure has also been used by Coriasso et al. [53] to make a switch operating at 1.55 μm wavelength. In that case the switch was a dual-channel contradirectional coupler in which the switching was controlled by another optical beam at approximately the same wavelength. Thus an all-optical switch was achieved. The required control energy was ~ 1 pJ.

18.4 Quantum-Well Detectors

The unique properties of MQW structures have been shown to be useful in several different types of photodetectors. In some of these devices photocurrent gains as high as 2×10^4 have been observed.

18.4.1 Photoconductive Detectors

An extremely large photocurrent amplification phenomenon has been observed at very low voltages in superlattices composed of 35-Å thick wells of $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ interleaved with 35-Å thick barriers of $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ grown by molecular beam epitaxy [54]. Responsivities as high as $4 \times 10^3 \text{ A/W}$ were measured at $1.3 \mu\text{m}$ wavelength with 1.4 V bias, corresponding to a current gain of 2×10^4 . Current gain of approximately 50 was observed for a bias voltage as low as 20 mV. This large gain arises from the large difference in tunneling rates of electrons and holes through the barriers because of their different effective masses in the MQW structure. In these semiconductors the electron effective mass is only about $0.05m_0$; thus, they have a long quantum wavelength and can easily tunnel through the 35 Å thick barriers. The hole effective mass, however, is about $0.5m_0$, so they have a much shorter wavelength and a much lower probability of tunneling. The substantial difference in the hole and electron tunneling rates results in a long electron lifetime and a large photocurrent gain. Because of the selective tunneling characteristics of the MQW structure, favoring the motion of electrons, these devices have been called effective mass filters (EMF). The EMF has inherently lower dark current and noise than a conventional photodiode because of the low voltage at which it can operate. Also, the EMF can be tuned for maximum gain by changing the thickness of the layers. If the thickness of the layers is adjusted to make it an integral multiple of the electron wavelength, resonant tunneling occurs and the gain is greatly increased. (The reported gain of 2×10^4 was for this resonant tunneling case).

Quantum photoconductivity has also been observed in a forward-biased p⁺-n junction with an $\text{Al}_{0.48}\text{In}_{0.52}\text{As}(23 \text{ \AA})/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}(49 \text{ \AA})$ superlattice in the n layer [55]. In that case a photocurrent gain of 2×10^3 was observed. Two other interesting features of the photoresponse were a large blue shift in the spectral response and a reversal of the direction of flow of the photocurrent when the magnitude of the forward bias voltage exceeded the built-in potential of the junction.

18.4.2 MQW Avalanche Photodiodes

$\text{Al}_{0.48}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$ multiquantum well structures have also been used to make $1.3 \mu\text{m}$ wavelength avalanche photodiodes [56]. A p⁺-i-n⁺ structure was grown by MBE, consisting of a 35-period $\text{Al}_{0.48}\text{In}_{0.52}\text{As}/\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$, 139-Å thick, MQW region sandwiched between p⁺ and n⁺ $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$ transparent layers.

DC current gain of 62 was measured. These devices were found to have good high frequency response, exhibiting an FWHM pulse response of 220 ps at a photocurrent gain of 12. The dark current at unity gain was 70 nA.

Low dark current InP/Ga_{0.47}In_{0.53}As superlattice avalanche photodiodes, operating at wavelengths in the range of 1.3 μm to 1.6 μm have also been reported [57]. These devices had a superlattice consisting of 20 periods with 300-Å wells and 300-Å barriers. A photocurrent gain of 20 was measured at 0.9 V reverse bias, with a dark current of less than 10 nA. MQW avalanche photodiodes have been studied in regard to doping effects [58], multiplication noise [59] and manufacturing yield [60].

18.5 Self-Electro-Optic Effect Devices

In Section 18.3.3, an optically switched MQW switch was described. That device employed the nonlinear electro-optic effect to produce a phase shift and resultant switching in a dual-channel coupler. However, it required an excessively large optical power (~ 1 mW) to produce switching. The dream of an optical computer requires low energy switching if it is to be realized. A device which offers promise of meeting that requirement is the self-electro-optic device (SEED) [61]. The fundamental concept of the SEED involves the combination of low power quantum well modulators with photodetectors to produce devices with both optical inputs and outputs. Actually, a single diode can function simultaneously as both modulator and detector. In this case a p-i-n MQW diode is reverse biased through a bias resistor and light of the proper wavelength is introduced in the direction normal to the quantum well layers. The MQW structure is fabricated so as to place the excitonic absorption peak at the operating wavelength with zero bias applied to the diode. Then the bias voltage is chosen so that the excitonic absorption peak is shifted away from the desired operating wavelength when the full bias voltage is applied. Thus with no light falling on the input face of the diode, no photocurrent flows through the resistor and the full bias voltage is applied to the diode. However, when light is present at the input, photocurrent is generated and part of the bias voltage is dropped across the bias resistor. The reduced voltage across the diode results in increased absorption as the excitonic peak shifts toward the operating wavelength. This positive feedback results in a “snowballing” effect which switches the diode to a high absorption, low voltage state when the input optical power exceeds a certain critical threshold level. If the input light level is reduced, the same positive feedback mechanism causes the diode to switch back to a low absorption, high voltage state when a certain lower sustaining level is passed. Thus the SEED is an optically switchable bistable device. An optical crossover switching network in which MQW SEED arrays are used as electrically addressed four-function interchange nodes has been produced by Cao et al. [62]. This network features a 64×64 matrix of high-speed interchange nodes.

By replacing the bias (load) resistor in the previously described circuit with a second identical MQW diode it is possible to make a symmetric SEED (S-SEED)

[63]. The key feature of the S-SEED is that it has gain, which permits switching at low levels of optical power. The S-SEED has two optical inputs, P_{in1} and P_{in2} , and two optical outputs, P_{out1} and P_{out2} , corresponding to light transmitted through each of the MQW diodes. If P_{in1} and P_{in2} are maintained equal the S-SEED will be in one state or another, with P_{out1} high and P_{out2} low, or vice versa, depending on the initial status of the diodes. If P_{in1} and P_{in2} are then both reduced by an identical amount the S-SEED will not switch. The switching of the device only occurs when the photocurrent in one diode starts to exceed that in the other one. In other words, the switching of the S-SEED is controlled by the ratio of input optical powers, not by their absolute levels. Thus the power in both input beams can be temporarily reduced, during which time an additional small amount of power added to one input beam can cause switching. Then the power of both beams can be increased. In this way a “time-sequential gain” is produced in which a very low power optical input signal can control a much larger output power. Switching energies of a few picojoules have been observed in S-SEEDs, and switching speeds of less than 1 ns have been demonstrated (at higher switching energies). Since these are relatively new devices, the prospects for further improvements in performance are good.

S-SEED arrays as large as 64×32 have been produced [64]. This array is only 1.3 mm square, with each S-SEED consisting of two MQW mesa diodes, $10 \mu\text{m} \times 10 \mu\text{m}$ in size, connected in series electrically. Only two electrical connections are required for the bias supply. In terms of the number of logic gates available, the 16×32 S-SEED array corresponds to a chip with 6144 logical pinouts, with two inputs and one output for each gate, although the chip actually has 16 384 physical “pinouts” because each gate has two optical input beams and two output beams. (A logical “1” state is one beam more powerful than the other; the “0” state is the reverse. Either $P_1 > P_2$ or $P_2 > P_1$ can be defined as the “1” state.)

S-SEEDs have been proposed for use in other applications, such as smart pixels, A/D converters, dynamic memories and oscillator circuits, as well as for digital logic [65–67]. It is too soon to say whether the S-SEED will evolve into the basic element of future optical computers, but it is definitely a potential candidate because of its small size and relatively low switching energy. Thus the SEED is continuing to be a device that is currently studied.

18.6 Quantum-Well Devices in OEIC’s

Quantum well devices have been incorporated into a number of different OEIC’s for a variety of applications. The fact that quantum well devices can be fabricated in the same III–V semiconductor materials that are preferred for other opto-electronic devices facilitates their monolithic integration. One example of such integration has already been presented in Section 18.3.3, in which a 2×2 MQW gate matrix switch was described. In the remaining paragraphs of this chapter additional examples will be given.

18.6.1 Integrated Laser/Modulators

Since one of the major applications envisioned for MQW devices is the modulation of semiconductor diode lasers, the monolithic integration of laser diodes with MQW modulators is highly desirable. Such integration could reduce the optical coupling losses between laser and modulator, as well as facilitating the introduction of the modulator electrical driving signal if a microwave stripline were deposited on the surface of the chip. Of course, monolithic integration should also ultimately reduce the cost and improve the reliability of the laser/modulator pair.

A monolithically integrated quantum well modulator and laser diode have been made in GaAs/GaAlAs by Tarucha and Okamoto [68]. Both the laser and the modulator had the same MQW structure, as shown in Fig. 18.8. The MQW structure consisted of 16 periods of 80-Å thick GaAs wells and 50-Å thick $\text{Ga}_{0.75}\text{Al}_{0.25}\text{As}$ barriers, sandwiched between 1.5 μm thick $\text{Ga}_{0.57}\text{Al}_{0.43}\text{As}$ cladding layers, in a 20- μm wide stripe waveguide mesa. The laser diode and modulator were separated by a 2.6- μm wide gap formed by reactive ion beam etching. The waveguide length was 110 μm for the laser and 190 μm for the modulator. The capacitance of the modulator was 7 pF for zero bias voltage, and the reverse breakdown voltage was 9 V. The laser threshold current was approximately 30 mA. With the laser diode biased at $1.4I_{\text{th}}$, a low frequency modulation depth of 7 dB was measured. The high-frequency modulation response curve was flat out to about 0.3 GHz then fell off to a -3 dB degradation at about 0.9 GHz, limited by the RC time constant of the modulator. Reducing the waveguide width should improve not only the modulator frequency response but also the laser threshold current. For 13- μm wide, singlemode waveguide the predicted capacitance is 0.4 pF, leading to a projected 10 GHz bandwidth.

An InGaAsP/InP DFB laser and an InGaAs/InAlAs MQW modulator have been monolithically integrated on an InP substrate by Kawamura et al. [69, 70]. The emission wavelength was 1.55 μm . Notably the DFB laser was grown by liquid phase epitaxy, while the MQW modulator was produced by molecular beam epitaxy. First the DFB structure, consisting of a 0.15- μm InGaAsP active layer and a 0.15- μm waveguiding layer (with DFB corrugations), a 2.4- μm InP cladding layer and a

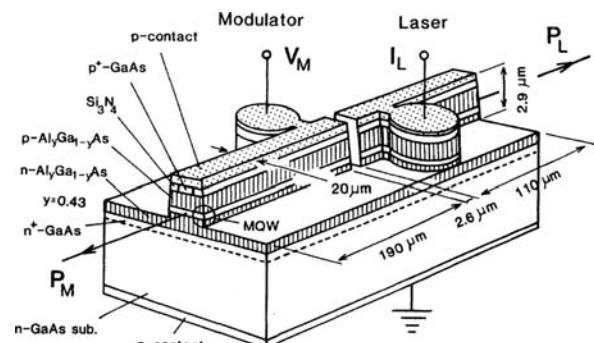


Fig. 18.8 Integrated MQW laser and modulator [68]

0.5- μm waveguiding layer (with DFB corrugations), a 2.4- μm InP cladding layer and a 0.5- μm InGaAsP cap, was grown by LPE on a Si-doped InP substrate. This DFB structure was then removed from the portion of the substrate on which the modulator was to be grown by using masked etching. The InGaAs/InAlAs MQW p-i-n modulator structure was then grown by MBE using SiO_2 masking. The MQW structure consisted of 40 periods of undoped InGaAs/InAlAs (65- \AA wells, 70- \AA barriers) sandwiched with a Si-doped n-InAlAs layer and Be-doped p-InAlAs layer and capped with a 0.5- μm InGaAs layer. Finally, the DFB laser was masked and etched to produce a 10- μm wide stripe geometry laser (cavity length = 300 μm and the MQW p-i-n layers were formed into a mesa-type optical modulator (mesa width 100 μm , cavity length 50 μm). The laser/modulator operated at 1.55 μm with a single longitudinal mode. An off-on ratio as high as 3 dB was observed, with a pulse response time of 300 ps.

Another example of an integrated MQW laser/modulator, of a different type than those described previously, is the monolithically integrated laser transmitter of Nakano et al. [71]. In that case, an MQW GaAlAs laser was integrated on a Semi-insulating GaAs substrate with a driver circuit containing 10 field effect transistors. The FETs were fabricated by selective ion implantation of dopant atoms into the semi-insulating substrate. The laser had a threshold current density of 40 mA, and could be modulated at a data rate of 2 Gbit/s using non-return-to-zero (NRZ) coding.

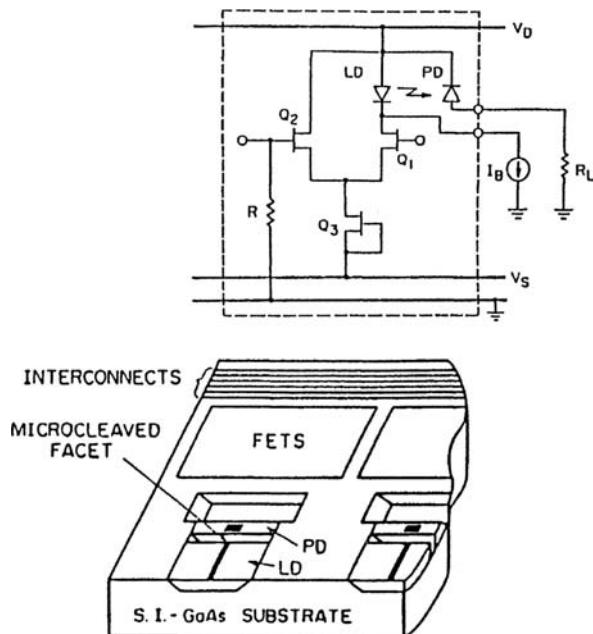
An example of an integrated MQW device intended for OEIC receiver circuits is the monolithically integrated InGaAs/AlGaAs pseudomorphic single-quantum-well MODFET and avalanche photodiode reported by Zebda et al. [72]. In this circuit a 1 μm gate length FET is combined with a 30 $\mu\text{m} \times$ 50 μm APD which contains a 13-period GaAs/AlGaAs (400/400 \AA) MQW structure in its avalanching region.

Clearly MQW lasers, modulators and detectors can be monolithically integrated with other opto-electronic devices and with conventional electronic devices as well. Most of the integration so far has involved only one laser and a modulator. However, more sophisticated circuits will surely evolve. The four-channel transmitter array described in the next section is an example of one such circuit.

18.6.2 A Four-Channel Transmitter Array with MQW Lasers

A monolithically integrated four-channel opto-electronic transmitter array has been fabricated by Wada et al. [73], as shown in Fig. 18.9. Each transmitter channel consists of a GaAlAs MQW laser diode, a monitor photodiode, three FET's and a resistor. All components are fabricated monolithically on a single semi-insulating GaAs substrate. The mesa-etched, n-channel FET's were designed to have a transconductance of greater than 17 (milliohms) $^{-1}$, and the resistors were made to be 50 Ω to match the input to an external microstrip line. The laser/photodiode pairs were aligned on the edge of the chip and an embedded heterostructure was used for both the lasers and monitor photodiodes. Etched wells were formed in the substrate to iso-

Fig. 18.9 Integrated four-channel transmitter [73]



late the laser, photodiode and FETs from one another. The laser facets were cleaved after a graded index, separate confinement heterostructure, single quantum well (GRIN-SCH-SQW) laser structure was grown on the substrate by MBE. A ridge waveguide, 3 μm wide, was formed by etching the laser surface prior to metallization. The final chip had dimensions of 4×2 mm, with 1-mm separation between the laser stripes. The laser cavity length was 40 μm , and the separation between the lasers and photodiodes was 60 μm .

Threshold currents for the lasers were measured to be in the range of 17.5 ± 3.5 mA, and the differential quantum efficiencies were $54.5 \pm 9.5\%$. The L - I curve for the lasers was linear up to an output level of 4 mW/facet. Light emission was essentially single mode, at a wavelength of 8340 Å. The photodiode response curve was also linear, with a dark current of less than 100 nA. The power dissipation of the array chip was estimated to be approximately 1.2 W. Modulation of the laser diodes of all four channels at a data rate of 1.5 Gbit/s (NRZ format) was possible. Crosstalk between channels was measured to be -28 dB at a frequency of 500 MHz, rising to a value of -14 dB at 1 GHz.

Preliminary aging tests on these lasers have determined an output power degradation rate of $8\%/\text{kh}$, after an initial “burn-in” period of 200 h. The overall electrical-optical conversion efficiency was measured to be 6 mW/V, which makes the transmitter suitable for operation with ECL inputs signals. Overall, considering the relatively high levels of integration in this OEIC, the performance characteristics of the various components (and of the complete chip) are quite good. They clearly

demonstrate the benefits obtainable by the monolithic integration of quantum well devices. Further improvements in technology will surely lead to even more impressive results.

In addition to the OEICs produced by monolithic integration, hybrid integration of MQW devices also is capable of yielding effective circuits. For example, Gruber [74] has reported a multi-chip module that incorporates arrays of MQW p-i-n photodiodes along with flip-chip-bonded vertical-cavity surface emitting laser arrays onto a fused-silica substrate in order to produce an OEIC with planar-integrated free-space optical vector-matrix-type interconnects.

Problems

- 18.1 Explain the difference between a multiple quantum well structure and a super-lattice.
- 18.2 For a deep quantum well of GaAs with a thickness $L_z = 70 \text{ \AA}$, $L_x = L_y = 100 \mu\text{m}$, calculate the magnitude of the steps in the conduction band cumulative density of states function and the energies at which they occur relative to the bottom of the well for the $n = 1, 2$ and 3 levels. (Assume that the well has infinitely high sidewalls.)
- 18.3 If the quantum well of Problem 18.2 were incorporated into a diode laser, what would be the emission wavelength for electron transitions between the $n = 1$ level in the conduction band and the $n = 1$ level in the valence band? (Assume 300 K operation.)
- 18.4 (a) Explain how a single p-i-n MQW diode can function as a self-electro-optic-effect device (SEED).
 (b) Why are SEEDs envisioned as being possible digital logic elements in a computer?
- 18.5 A single-quantum well of GaAs with a thickness of 60 \AA is incorporated into a GaAlAs heterojunction laser. The laser has a stripe geometry, being $200 \mu\text{m}$ long and $10 \mu\text{m}$ wide.
 - (a) What is the emission wavelength if the energy bandgap is $E_g = 1.52 \text{ eV}$ in the emitting region and electron transitions occur between the $n = 1$ level in the conduction band and the $n = 1$ level in the valence band?
 - (b) If the gain coefficient of the laser in part (a) is $g = 80 \text{ cm}^{-1}$, what would be the gain coefficient of a multiple quantum well laser with 20 identical wells?
- 18.6 A certain Quantum-Confining-Stark-Effect modulator has a geometry such that the light passes through it in a direction perpendicular to the plane of the p-n junction. With no voltage applied the absorption coefficient at the operating wavelength is 10 cm^{-1} and with the maximum allowable bias voltage applied the absorption coefficient is $2 \times 10^3 \text{ cm}^{-1}$. The thickness of the p-n junction (active region) is $10 \mu\text{m}$.

- (a) What is the on/off ratio of the modulator?
- (b) If the same device structure were fabricated so as to have the light pass through it parallel to the junction, with a confinement factor of 0.8, what would be the new on/off ratio if the device were $50 \mu\text{m}$ long?
- 18.7 For a deep quantum well of $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ with a thickness $L_z = 70\text{A}$, $L_x = L_y = 100 \mu\text{m}$, calculate the magnitude of the steps in the conduction band cumulative density of states function and the energies at which they occur relative to the bottom of the well for the $n = 1, 2$ and 3 levels. Assume that the well has infinitely high sidewalls. For $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ the effective masses of electrons and holes are $0.045m_0$ and $0.465m_0$, respectively.
- 18.8 A MQW laser has a threshold current of 10 mA at a temperature of 21°C . If its threshold temperature coefficient $T_0 = 425^\circ\text{C}$, what will be its threshold current at a temperature of 45°C .
- 18.9 Why do MQW lasers have a narrower emission linewidth than that of conventional heterojunction lasers without a MQW structure?

References

1. K. Seeger: *Semiconductor Physics*, 5th edn., Springer Ser. Solid-State Sci., Vol. 40 (Springer, Berlin, Heidelberg 1991)
2. R. Dingle: *Proc. 13th Int'l Conf. on Physics of Semiconductors*, F.G. Fumi (ed.), (North-Holland, Amsterdam 1976)
3. A. Yariv: *Quantum Electronics*, 3rd edn. (Wiley York, New York 1988) pp. 267–268
4. A. Yariv: *Quantum Electronics*, 3rd edn. (Wiley York, New York 1988) p. 270
5. A. Yariv: *Quantum Electronics*, 3rd edn. (Wiley York, New York 1988) p. 272
6. A. Larsson, P. Andrekson, B. Jonsson, C. Lindstrom: IEEE J. **QE-25**, 2013 (1989)
7. J. Feldmann, G. Peter, E.O. Göbel, K. Leo, H. Polland, K. Ploog, K. Fujiware, T. Nakayama: Appl. Phys. Lett. **54**, 226 (1987)
8. W.T. Tsang: Appl. Phys. Lett. **40**, 217 (1982)
9. L.M. Walpita: J. Opt. Soc. Am. A **2**, 592 (1985)
10. PL. Derry, A. Yariv, K.Y. Lau, N. Bar-Chaim, K. Lee, J. Rosenberg: Appl. Phys. Lett. **50**, 1773 (1987)
11. W.J. Choi, P.D. Dapkus: Self-defined oxide-current-aperture buried-heterostructure ridge waveguide InGaAs single-quantum-well diode laser. IEEE Photonics Tech. Lett. **11**, 773 (1999)
12. W.T. Tsang: Appl. Phys. Lett. **39**, 786 (1981)
13. R. Chin, N. Hollonyak Jr., B. Vojak, K. Hess, R. Dupuis, P. Dapkus: Appl. Phys. Lett. **36**, 19 (1980)
14. Z.Y. Yu, V. Kreismanis, C.L. Tung: Appl. Phys. Lett. **44**, 136 (1984)
15. Y. Sasai, J. Ohya, M. Ogura, T. Kajiwara: Electron. Lett. **23**, 232 (1987)
16. Y. Sasai, N. Huse, M. Ogura, T. Kajiwara: J. Appl. Phys. **59**, 28 (1986)
17. Y. Sasai, M. Ogura, T. Kajiwara: J. Cryst. Growth **78**, 461 (1986)
18. Y. Arakawa, A. Yariv: IEEE J. **QE-21**, 1666 (1985)
19. M. Okai, T. Tsuchiya, A. Takai, N. Chinone: IEEE Photon. Tech. Lett. **4**, 526 (1992)
20. J.J. Coleman: Strained-layer InGaAs quantum well heterostructure lasers. IEEE J. Selected Topics Quantum Electron. **6**, 1008 (2000)
21. A. Cavicchia, O. Lang, D. Gershoni, A. Sergent, J. Vandenberg, S.N.G. Chu, M.B. Panish: Appl. Phys. Lett. **54**, 739 (1989)

22. P.J.A. Thijs, T. van Dongen, B.H. Verbeek: OFC'90, San Francisco, Paper WJ2
23. J.J. Coleman: OFC'90, San Francisco, Paper WJI
24. P-H. Leia, M-Y. Wu, C-C. Linb, W-J. Hob, M-C. Wu: High-power and low-threshold-current operation of 1.3 μm strain-compensated AlGaInAs/AlGaInAs multiple-quantum-well laser diodes, Solid-State Electron. **46**, 2041 (2002)
25. Z. Alferov: Double heterostructure lasers: early days and future perspectives. IEEE J. Selected Topics Quant. Elect. **6**, 832 (2000)
26. Y. Arakawa, K. Vahala, A. Yariv: Appl. Phys. Lett. **45**, 950 (1984)
27. H. Temkin, G. Dolan, M.B. Panish, S.N.G. Chu: Appl. Phys. Lett. **50**, 413 (1987)
28. J. Cibert, P.M. Petroff, G.J. Dolan, C.J. Pearton, A.C. Gossard, J.H. English: Appl. Phys. Lett. **49**, 1275 (1986)
29. G. Park, D.L. Huffaker, Z. Zou, O.B. Shchekin, D.G. Deppe: Temperature dependence of lasing characteristics for long-wavelength (1.3- μm) GaAs-based quantumdot lasers. IEEE Photonics Tech. Lett. **11**, 301 (1999)
30. W. Tsang, M. Wu, Y. Chen, F. Chou, R. Logan, S. Chu, A. Sergant, P. Magill, K. Reichmann, C. Burrus: IEEE J. **QE-30**, 1370 (1994)
31. D.A.B. Miller, D.S. Chemia, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood, L.A. Burrus: Phys. Rev. Lett. **53**, 2173 (1984)
32. D.S. Chemia, T.C. Damen, D.A.B. Miller, A.C. Gossard, W. Wiegmann: Appl. Phys. Lett. **42**, 864 (1983)
33. T.H. Wood, C.A. Burrus, D.A.B. Miller, D.S. Chemia, T.C. Damen, A.C. Gossard, W. Wiegmann: IEEE J. **QE-21**, 117 (1985)
34. D.A.B. Miller, D.S. Chemia, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood, C.A. Burrus: Phys. Rev. B **32**, 1043 (1985)
35. T.H. Wood: IEEE J. LT-**6**, 743 (1988)
36. P.J. Stevens, G. Parry: IEEE J. LT-**7**, 1101 (1989)
37. T.H. Wood, C.A. Burrus, R.S. Tucker, J.S. Weiner, D.A.B. Miller, D.S. Chemia, T.C. Damen, A.C. Gossard, Wiegmann: Electron. Lett. **21**, 693 (1985)
38. T.H. Wood, R.W. Tkach, A.R. Chraplyvy: Appl. Phys. Lett. **50**, 798 (1987)
39. T.H. Wood: Appl. Phys. Lett. **48**, 1413 (1986)
40. K. Wakita, K. Sato, I. Kotaka, M. Yamamoto, T. Kataoka: Electron. Lett. **30**, 302 (1994)
41. P.S. Cho, D. Mahgerefteh, J. Coldhar: All-optical 2R regeneration and wavelength conversion at 20 Gb/s using an electroabsorption modulator. IEEE Photonics Tech. Lett. **11**, 1662 (1999)
42. T. Ido, S. Tanaka, M. Koizumi, H. Inoue: Ultrahigh-speed MQW electroabsorption modulators with integrated waveguides, Proceedings of SPIE **3006** Optoelectronic Integrated Circuits, Y-S. Park, R.V. Ramaswamy, (eds.), (1997) pp. 282–290
43. T. Kuri, K. Kitayama, A. Stohr, Y. Ogawa: Fiber-optic millimeter-wave downlink system using 60 GHz-band external modulation. IEEE J. Lightwave Tech. **17**, 799 (1999)
44. J.I. Pankove: *Optical Processes in Semiconductors* (Prentice-Hall, Englewood Cliffs, NJ 1971) pp. 89–90
45. T. Hiroshima: Appl. Phys. Lett. **50**, 968 (1987)
46. H. Nagai, M. Yamanishi, Y. Kam, I. Suemune: Electron. Lett. **22**, 888 (1986)
47. T.H. Wood, R.W. Tkach, A.R. Chraplyvy: Appl. Phys. Lett. **50**, 798 (1987)
48. I.D. Hennings, J.V. Collins: Electron. Lett. **19**, 927 (1983)
49. J.S. Weiner, D.A.B. Miller, D.S. Chemia: Appl. Phys. Lett. **50**, 842 (1987)
50. P. Li Kam Wa, J.E. Sitch, N.J. Mason, J.S. Roberts, P.N. Robson: Electron. Lett. **21**, 26 (1985)
51. A. Ajisawa, M. Fujiware, J. Shimizu, M. Sugimoto, M. Uchida, Y. Ohta: Electron. Lett. **23**, 1121 (1987)
52. T.H. Wood, J.Z. Pastalan, C.A. Burrus, B.I. Miller, J.L. de Miguel, U. Koren, M. Young: OSA/IEEE Conf. on Integrated Photonics Research, Hilton Head, SC (1990) Paper TuG4
53. C. Coriasso, D. Campi, L. Faustini, A. Stano, C. Cacciatore: Optically controlled contradirectional coupler. IEEE J. Quantum Electron. **35**, 298 (1999)

54. F. Capasso, K. Mohammed, A.Y. Cho, R. Hull, L. Hutchinson: *Appl. Phys. Lett.* **47**, 420 (1995)
55. F. Capasso, K. Mohammed, A.Y. Cho, R. Hull, L. Hutchinson: *Phys. Rev. Lett.* **55**, 1152 (1985)
56. K. Mohammed, F. Capasso, J. Allam, A.Y. Cho, L. Hutchinson: *Appl. Phys. Lett.* **47**, 597 (1985)
57. F. Capasso: OFC/IGWO'86m Atlanta, GA, Paper WCCI
58. I. Yun, H.M. Menkarl, Y. Wang, I.H. Oguzman, J. Kolnik, K.F. Brennan, G.S. May, C.J. Summers, B.K. Wagner: Effect of doping on the reliability of GaAs multiple quantum well avalanche photodiodes. *IEEE Trans. Electron. Devices* **44**, 535 (1997)
59. A. Salokalve, M. Toivonett, M. Hovinen: Multiplication noise in GaAs/AlGaAs multiquantum well avalanche photodiodes with different well different well widths. *Electron. Lett.* **28**, 416 (1992)
60. I. Yun, G.S. May: Parametric manufacturing yield modeling of GaAs/AlGaAs multiple quantum well avalanche photodiodes. *IEE Trans. Semiconductor Manufacturing* **12**, 238 (1999)
61. For a review of SEED devices see, e.g., D.A.B. Miller: *Optics and Photon. News* **1**, 7 (April 1990) or H.S. Hinton, A.L. Lentine: Multiple quantum-well technology takes SEED, *IEEE Circuits and Devices Magazine* **9**, 12 (1993)
62. M. Cao, H. Li, A. Jun, F. Luo, X. Jun, L. Nu, W. Gao: *Opt. Laser Tech.* **26**, 271 (1994)
63. A.L. Lentine, H.S. Hinton, D.A.B. Miller, J.E. Henry, J.E. Cunningham, L.M.F. Chirovsky: *IEEE J. QE-25*, 1928 (1989)
64. L.M.F. Chirovsky, L.A. D'Asaro, R.F. Kopf, J.M. Kuo, A.L. Lentine, F.B. McCormick, R.A. Novotny, G.D. Boyd: OSA Annual Meeting, Orlando, FL (1989) Paper PD28
65. H.D. Chen, K. Liang, Q.M. Zeng, X.J. Li, Z.B. Chen, Y. Du, R.H. Wu: Flip-chip bonded hybrid CMOS/SEED optoelectronic smart pixels. *IEEE Proceedings, Optoelectronics* **147**, 2 (2000)
66. M. Moran, C.J. Rees, J. Woodhead: Operating characteristics of GaAs-InGaAs selfbiased piezoelectric S-SEEDs: *IEE Proc., Opt.* **146**, 31 (1999)
67. S.F. Al-Sarawi, P.B. Atanackovic, W. Marwood, B.A. Clare, K.A. Corbett, K.J. Grant, J. Munch: Differential oversampling data converters in SEED technology, *Microelectron. J.* **33**, 141 (2002)
68. S. Tarucha, H. Okamoto: *Appl. Phys. Lett.* **48**, 1 (1986)
69. Y. Kawamura, K. Wakita, Y.I. Taya, Y. Yoshikuni, H. Asabi: *Electron. Lett.* **22**, 242 (1986)
70. Y. Kawamura, K. Wakita, Y. Yoshikuni, Y. Itaya, H. Asahi: *IEEE J. QE-23*, 915 (1987)
71. H. Nakano, S. Yamashita, T. Tanaka, H. Hirao, N. Naeda: *IEEE J. LT-4*, 574 (1986)
72. Y. Zebda, R. Lipa, M. Tutt, D. Pavlidis, P. Bhattacharya, J. Pamulapati, J. Oh: *IEEE Trans. ED-35*, 2435 (1988)
73. O. Wada, N. Nobuhara, T. Sanada, M. Kuno, M. Makiuchi, T. Fujii, T. Sakurai: *IEEE J. LT-7*, 186 (1989)
74. M. Gruber: Multichip module with planar-integrated free-space optical vector-matrix-type interconnects, *Appl. Opt.* **43**, 463 (2004)

Supplementary Reading on Quantum Wells

- P. Bhattacharya: *Semiconductor Optoelectronic Devices* (Prentice-Hall, Englewood Cliffs, NJ (1994) pp. 133–137, 294–299
- D.A.B. Miller, D.S. Chemia, S. Schmitt-Rink: Electric field dependence of optical properties of semiconductor quantum wells, in *Semiconductors*, H. Haug (ed.) (Academic, New York 1988) pp. 325–360

Chapter 19

Micro-Optical-Electro-Mechanical Devices

In recent years techniques of micromachining, such as ion-beam sputtering, plasma sputtering and selective chemical etching, have been developed to the point that it is now possible to fabricate mechanical microstructures monolithically on the same chip as electronic and optical devices. This has lead to a new class of integrated devices or circuits, which are called micro-opto-electro-mechanical devices or systems (MOEMS). Since this acronym is somewhat unwieldy, and since micro-electro-mechanical devices (MEMs) were developed before those including optical elements, the latter acronym is often used for both types of devices. Quite a few different types of MOEMS or MEMs have been fabricated to perform a wide variety of functions. Generally they can be divided into three different classes, sensors, actuators and optical elements. Examples of devices in each class are given in Table 19.1.

The distinction between the three classes is somewhat indefinite, and no generally accepted classification exists. However it seems reasonable to say that a sensor enables one to detect or measure some property, an actuator has moving parts, and an optical element performs its function without moving parts. Detailed descriptions of some of the MEMs listed in Table 19.1 will be given later in this chapter; but first, it is necessary to review some of the basic equations of mechanics that electrical and optical engineers may not be accustomed to using.

Table 19.1 Micro-opto-electro-mechanical systems

Sensors	Actuators	Optical elements
Length	Tuner	Microlens
Pressure	Switch	Coupler
Vibration	Scanner	Filter
Light	Interferometer	Micromirror
Voltage	Voltage Generator	Beam Splitter
Temperature	Micromotor	HOE (grating)

19.1 Basic Equations of Mechanics

Mechanics is the field of engineering science that deals with forces and the effects that they have on various materials. To give one a full understanding of the field is elements beyond the scope of this book. However, a brief review of the key design equations will suffice to permit one to understand how MEMs operate and to design new MEMs. For a thorough discussion of the principles of mechanics see, for example, Popov [1].

19.1.1 Axial Stress and Strain

Assume a solid bar of initial length = L_0 and diameter = D . If a tensile force = F is uniformly applied to the ends of the bar, the bar will lengthen by an amount = ΔL , as shown in Fig. 19.1.

In that case the axial strain ε_a is given by

$$\varepsilon_a = \frac{\Delta L}{L_0}. \quad (19.1)$$

The stress is given by

$$\sigma = \frac{F}{\text{Area}} = \frac{F}{\pi(D/2)^2}. \quad (19.2)$$

Tensile stresses are normally designated to be negative while compressive stresses are taken to be positive. If the stress vs. strain curve of a material is plotted on a linear scale, the curve is a straight line up to a break point (approximately the yield strength) at which the curve becomes sublinear. Stress and strain are related by Hooke's Law up to this break point

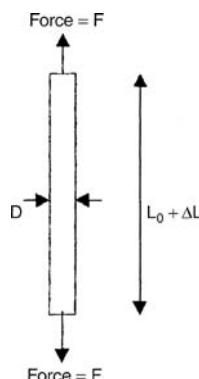


Fig. 19.1 Bar subjected to axial force

$$\sigma = \varepsilon E, \quad (19.3)$$

where E , the slope of the linear portion of the stress-strain curve, is called Young's modulus. Some materials are brittle and have a sharp break point and others are ductile, having a gradual departure from linearity. In either case the maximum value of stress that is reached as strain is increased is approximately the yield strength. Stress (and yield strength) have units of force per unit area, while strain is a unitless quantity. Thus, Young's modulus also has units of force per unit area, and is usually given in gigapascals or GPa. Yield strength and Young's modulus are important material properties that have been measured and published for most materials of interest [2].

Up to this point, all of the stress and strain considered has been axial, but shear stress and strain, caused by a force directed perpendicularly to the axis of the rod also can be important. Equations (19.1), (19.2) and (19.3) also hold for sheer stress and strain. The ratio of lateral strain to axial strain is defined as Poisson's ratio, given by

$$\nu = -\frac{\varepsilon_l}{\varepsilon_a}. \quad (19.4)$$

Defined in this way Poisson's ratio is always positive, since a tensional force that makes ε_a negative causes a constriction of the diameter of the rod that makes ε_l positive.

19.1.2 Thin Membranes

A number of MEMs, such as pressure and vibration sensors make use of thin membranes of semiconductors or other materials. The membranes are typically supported at only their periphery. Thus they are free to deflect over their surface in response to any force or pressure (force per unit area) that is applied. The deflection can then be used to measure the force or pressure. If the force or pressure is applied through electrical or magnetic means, the membrane acts as a transducer, converting electrical or magnetic energy into mechanical motion. This motion of the membrane can be used to move a reflecting surface so as to cause phase shift or modulation of an optical beam.

The equations relating the important parameters of thin membranes are as follows [3]:

Assume a square membrane with side length = a , thickness = t , Young's modulus = E , density = ρ , and Poisson's ratio = ν . If this membrane is exposed uniformly over its top surface to a pressure = P , its maximum deflection w_{\max} (occurring at the center) is given by

$$w_{\max} = 0.001265 Pa^4/D, \quad (19.5)$$

where D is the plate flexural rigidity, given by

$$D = \frac{Et^3}{12(1 - \nu^2)}. \quad (19.6)$$

Its maximum longitudinal stress σ_1 is given by

$$\sigma_t = 0.3081 P(a/t)^2. \quad (19.7)$$

Its maximum transverse stress σ_1 is given by

$$\sigma_t = \nu \sigma_l. \quad (19.8)$$

Its resonant frequency F_0 is given by

$$F_0 = (1.654t/a^2) \left[\frac{E}{\rho(1 - \nu^2)} \right]^{1/2}. \quad (19.9)$$

19.1.3 Cantilever Beams

Many MOEMs make use of a cantilever beam structure, a bar of material that is firmly held at one end but free to deflect at the other end. A cantilever reacts to a force, either concentrated at one point or distributed, by deflecting. The deflection depends on the magnitude of the force and on the nature of its distribution. Thus cantilevers can be used to measure forces. If the force is applied through electrical or magnetic means, the cantilever serves as a transducer, converting electrical energy into mechanical motion.

Assume a cantilever beam, fixed at one end ($x = 0$), of length = L , width = a , thickness = t , Young's modulus = E and density = ρ , as shown in Fig. 19.2.

If the beam is loaded with a uniformly distributed force per unit width $P = F/a$ over its surface, the deflection w over the length of the beam ($0 < x < L$) is given by [3]

$$W(P, x) = \frac{Px^2}{24EI}(6L^2 - 4Lx + x^2), \quad (19.10)$$

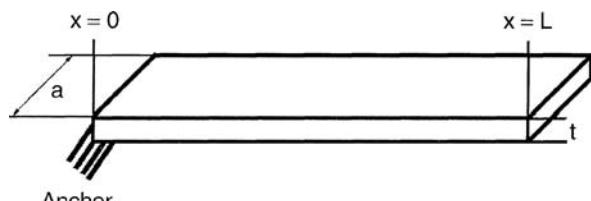


Fig. 19.2 Cantilever beam configuration

where I is the bending moment of inertia, given by

$$I = \frac{at^3}{12}. \quad (19.11)$$

The maximum stress σ_{\max} is given by

$$\sigma_{\max} = \frac{PL^2t}{4I}. \quad (19.12)$$

Alternatively, if the beam is loaded with a point load Q (where Q is a force) at the end of the beam ($x = L$), the deflection w is given by

$$W(Q, x) = \frac{Qx^2}{6EI}(3L - x). \quad (19.13)$$

and the maximum stress is

$$\sigma_{\max} = \frac{QLt}{2I}. \quad (19.14)$$

The frequency of the fundamental vibrational resonant mode of the beam F_0 is given by

$$F_0 = 0.161 \frac{t}{L^2} \left(\frac{E}{\rho} \right)^{1/2}. \quad (19.15)$$

19.1.4 Torsion Plates

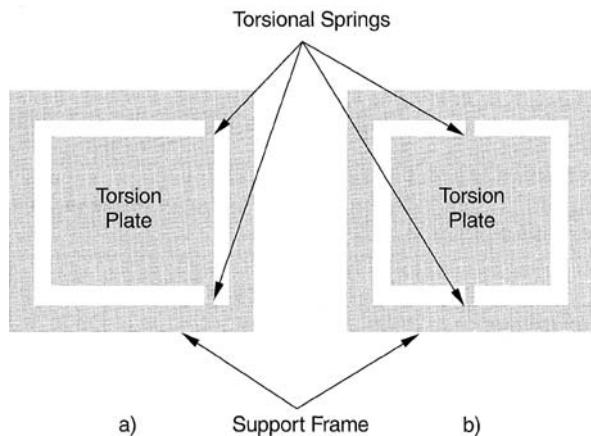
Thin plates of semiconducting material, supported only by small beams at two points can be fabricated by micro-etching techniques. Two typical configurations are shown in Fig. 19.3. Somewhat surprisingly, even in the case of a brittle semiconductor like silicon, the small beams supporting the plate can form torsional springs that permit the plate to rotate through an angle without snapping off [4, 5]. The torsion plate can be made to rotate by the application of either magnetic or electrical means, as will be discussed later in this chapter. Thus, such MOEMs as scanning mirrors and optical switches can be made.

As the plate rotates through an angle θ , the torsional springs develop a restoring torque T_r that is given by [4]

$$T_r = \frac{2(G\omega t^3\theta)}{3l} \left(1 - \frac{192}{\pi^5} \right) \frac{t}{\omega} \tan h \left(\frac{\pi\omega}{2t} \right), \quad (19.16)$$

where l , w and t are the length, width and thickness of the beam and G is the shear modulus of the semiconductor (or other material) out of which the beam is formed.

Fig. 19.3 Micromachined torsion plates (a) asymmetric, supported at one edge (b) symmetric, supported at center



19.2 Thin Membrane Devices

A number of different types of MOEMs can be made that employ a thin membrane as the key element. Deflection of the membrane can be used in a sensor to detect and measure force or pressure. For example, Zhoе et al. [6] has used a thin circular membrane of silicon to make an optically interrogated pressure sensor, as shown in Fig. 19.4.

The MOEM was formed by etching a shallow cavity, of depth $h = 0.53 \mu\text{m}$, and diameter $d = 600 \mu\text{m}$, in a glass substrate and then bonding a layer of silicon onto the glass over the cavity. Finally the silicon was etched to a thickness $t = 26 \mu\text{m}$ to form a membrane. In use, the device is optically interrogated through an optical fiber at a wavelength of $\lambda_0 = 850 \text{ nm}$, and the deflection is interferometrically measured. In this case of a circular membrane, the relationship between the deflection w_r at a point that is located at a distance r from the center of the membrane is given by [19.6]

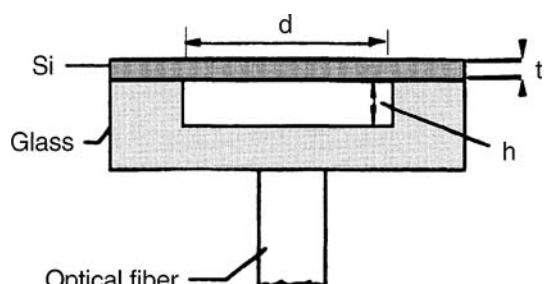


Fig. 19.4 Optically interrogated pressure sensor [6]

$$w_r = w_c \left[1 - \left(\frac{2r}{d} \right)^2 \right]^2,$$

where w_c is the deflection at the center of the membrane, which is given by

$$w_c = 3Pd^4 \frac{(1 - \nu^2)}{265Et^3}, \quad (19.18)$$

where ν is Poisson's ratio and E is Young's modulus of the material, in this case, silicon. (For crystalline silicon, $E = 160$ GPa, $\nu = 0.271$, and density $\rho = 2330$ kg/m³.) The device with the dimensions that have been given was designed to operate over a pressure range of 0 to 30 psi. At the maximum pressure, the deflection at the center was equivalent to $\lambda_0/8$.

A different type of membrane-based pressure sensor has been reported by Porte et al. [7]. Their device is a Mach-Zehnder interferometric pressure sensor designed to work as a remote sensor in a coherence modulation scheme, allowing a linear phase read-out of the signal. The interferometer was constructed from silicon nitride waveguides on a silicon substrate. Three membranes of thickness $t = 21$ μm were micromachined in cascade, centered under the reference arm of the device. In operation, a superluminescent diode emitting at $\lambda_0 = 1300$ nm is used as the light source. The relatively short coherence length of this source guarantees that the optical group path delay introduced by the interferometer is greater than the coherence length of the source. Hence, no interference can be detected directly at the output as a pressure variation is applied to the membranes, and the information is coded into the variation of the optical delay, and not into the transmitted output power, which remains constant. At the receiver, the decoding device is another similar interferometer with approximately the same optical group path delay. In this case, the intensity I at the output of the decoding stage is related to the pressure P applied to the membranes of the transmitter by [7]

$$I(P) = \frac{I_0}{2} \left[1 + \frac{1}{2} \cos\{\delta_2 - \delta_1 - \Delta_1(P)\} \right], \quad (19.19)$$

where I_0 is the source intensity, δ_1 and δ_2 are the optical delays of the transmitter and receiver Mach-Zehnder interferometers respectively and $\Delta_1(P)$ is the change in path delay produced by the pressure that is being sensed at the transmitter.

Silicon membranes can also be used to make deformable mirrors (DMs) that can be used to correct aberrations in optical beams. Since the membrane shape can be rapidly changed they can even be used to correct time-varying aberrations in imaging or beam forming applications. An example of this type of application has been reported by Bifano et al. [8]. In their device the mirror membrane was supported by an underlying membrane array of electrostatic parallel-plate actuators capable of deforming the flexible silicon mirror membrane. The three-layer structure, consisting of mirror membrane, actuator membrane and electrode layer, was fabricated on a

silicon substrate by surface micromachining and deposition of polysilicon layers in a multiuser foundry process [9]. Results from prototype devices demonstrated the feasibility of continuous and segmented-mirror MEMS-DM's with independently controllable zones. These mirrors exhibited ample stroke ($2 \mu\text{m}$) for compensation of typical aberrations in visible light. The position repeatability resolution of the system was approximately 10 nm. No hysteresis or temporal drift was observed in any devices, and yield of 95% was routinely achieved in prototype devices. The mirror bandwidth was measured to be dc-3.5 kHz. The electrostatic actuators operated without hysteresis and required about 200 V for operation. The entire package for the prototype mirror was approximately $35 \times 35 \times 6$ mm with an active mirror area of 7×7 mm for a 400-element MEMS-DM.

A moveable membrane of silicon can also be used to make an optical wave guide switch, for example as demonstrated by Veldhuis et al. [10]. They made an absorptive waveguide switch by locating a rectangular, moveable silicon membrane over the surface of a Si_3N_4 waveguide. The waveguide structure consisted of a ridged Si_3N_4 guide on top of a SiO_2 buffer layer and Si substrate. To produce the “off” state of the switch the membrane is moved close to the waveguide surface so that the evanescent tail of the waveguide mode extends into the silicon and strong absorption occurs at the operating wavelength $\lambda_0 = 632.8$ nm. For the “on” state the silicon membrane is moved away from the waveguide surface. The membrane structure can be either that of a bridge, clamped at both ends, or a cantilevered beam, clamped at one end. The motion is accomplished electrostatically by applying a voltage between the optical waveguide and the membrane. In one particular embodiment of the switch the silicon bridge memberane thickness was $20 \mu\text{m}$, the width was $120 \mu\text{m}$ and the length was 9.5 mm. For this device, the activation voltage required to pull the membrane in to its closest spacing to the waveguide (100 nm) was 2.5 V. (The mem-
brane is prevented from actually touching the waveguide surface by a series of small bumps that are fabricated in rows along the sides of the waveguide for that purpose.) When the switch was in this “off” position an extinction of 65 dB was observed. The fundamental resonance frequency of the bridge structure was calculated to be 1.75 kHz.

If the absorptive silicon bridge is replaced by a waveguiding bridge membrane, evanescent coupling between the waveguide on the substrate and the bridge waveguide can be switched on and off by electrostatically moving the bridge waveguide. Thus switching of the lightwaves between two output waveguide ports can be accomplished. A device of this type has been described by Cholett et al. [11].

As a final example of MOEMs that are based on a micromachined membrane, consider the reflective duplexer for fiber optic communication that has been described by Marxer et al. [12]. This device makes use of a Fabry-Perot reflective etalon, one mirror of which is formed on a moveable, 470 nm thick, polysilicon membrane. The membrane is deflected electrostatically to switch the reflectivity of the etalon at the operating wavelength. In this duplexer, both the input and output light travel through a single mode fiber that enters normal to the fabry-Perot reflecting surfaces. Behind the Fabry-Perot etalon is an InGaAs photodiode that detects the optical pulses of the downstream data. The etalon functions as a switchable mirror to

modulate the upstream data, while it passes the downstream data to the photodiode. For full duplex communication the upstream and downstream signals are separated using subcarrier modulation. While the upstream signal is base band modulated, a high frequency carrier is used for the down stream channel. The modulation rate for the upstream data is limited by the mechanical response of the membrane to approximately 2.5 Mb/s, while the detectable downstream data rate depends only on the response of the photodiode, which is several GB/s. A duplexer of this type, designed for operation at a wavelength of 1310 nm, had an insertion loss < 3 dB, a modulation voltage of 7 V_{pp} and a maximum modulation rate of 2.88 MB/s.

19.3 Cantilever Beam Devices

The evanescent-wave absorptive switch [10] that was mentioned in Section 19.2 was actually the first example of a cantilever beam device, since it was mentioned that it could be made clamped only at one end as well as clamped at both ends. Many other MOEMS based on a cantilever beam also have been demonstrated. One of the more interesting examples is the tunable micromachined vertical cavity laser described by Vail et al. [13]. This device uses a movable external mirror, suspended at the end of a cantilever beam, to change the length of the laser cavity, and hence to tune the emitted wavelength. The laser structure, which was fabricated in the GaAs/GaAlAs system is shown in Fig. 19.5.

The structure consists of a bottom DBR mirror, a quantum well active region, and a top DBR mirror that is suspended 1.2 μm above the semiconductor surface on a cantilever beam. When a voltage (charge) is applied to the top cantilever it induces an opposing polarity charge at the semiconductor surface and is electrostatically attracted toward it. Application of a voltage of about 5 volts was sufficient to produce a maximum deflection of 0.4 μm and a tuning range of 15 nm. The

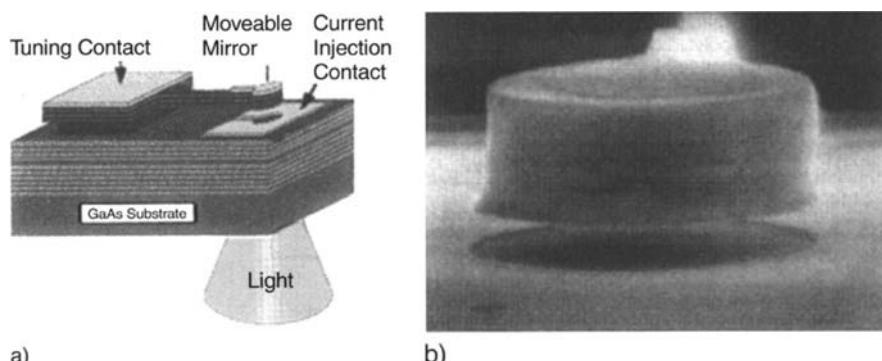


Fig. 19.5 a,b Tunable micromachined vertical cavity laser **a** diagram **b** microphotograph of moveable mirror [13] © 1995 IEEE

tuning power required was on the order of 250 pW. Since these VCSELs have a round mode shape and can be produced in a surface emitting array for convenient coupling to an optical fiber ribbon or bundle, it is expected that they will be particularly useful in optical communications systems. The authors also report having made tunable detectors [14] and filters [15] using the cantilever beam/suspended mirror approach. The filters had a tuning range of 70 nm with a bandwidth of 6 nm and 20.3 dB extinction ratio. The detectors had a tuning range of 30 nm, a bandwidth of 7 nm and an extinction ratio of 17 dB. For a review of the development of wavelength-tunable VCSELs and related structures, see Chang-Hasnain [16].

As might be expected, cantilever beam MOEMs can be conveniently used in switching applications. Chen et al. [17] have made a high-speed 2×2 optical switch by using a stress-induced, curved polysilicon cantilever beam to move a 45° angled mirror in and out of the central region between four optical fibers arranged in a pattern around the mirror, spaced at 90° angles to each other, as shown in Fig. 19.6.

The curvature of the polysilicon beam serves to move the mirror out of the optical path with no voltage applied to it. The upward curve was produced by depositing a stressed layer of Cr-Au onto the polysilicon beam. The beam curves upward because the polysilicon is under compressive residual stress while the Cr-Au is under tensile residual stress. To move the mirror into the optical path (a total vertical movement of 306 μm) a voltage is applied to the cantilever beam and electrostatic attraction draws it back into the lowered position, thus changing the switch from the transmission state to the reflection state. The pull down voltage was measured to be 20 V. The insertion loss of the switch was < 0.55 dB for the transmitting state and 0.7 dB for the reflecting state. Wavelength dependence throughout the range from 1500 to

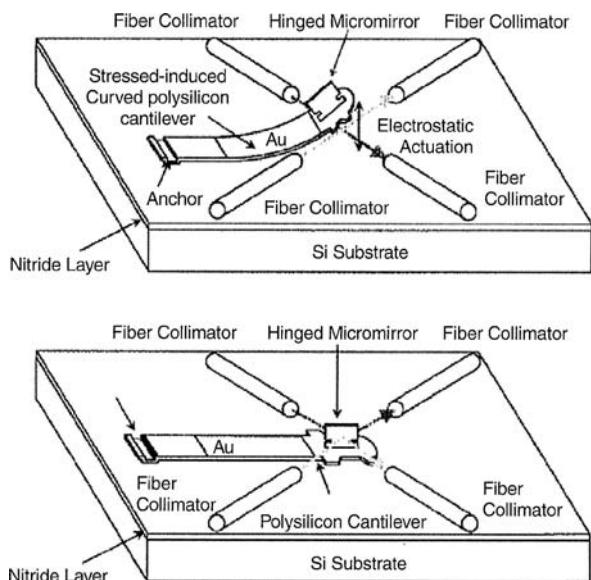


Fig. 19.6 Cantilever MOEM optical switch [17] © 1999 IEEE

1580 nm was < 0.12 dB, and polarization dependence over the range from 0 to 90° was < 0.09 dB. When a 20 V_{pp} square wave was used to toggle the switch, the turn-on time (corresponding to snap-down of the mirror) was 600 μ s. The turn-off time (corresponding to release of the mirror) was 400 μ s.

A cantilever-beam-based bistable switch also has been made in silicon by Hoffmann et al. [18]. In that case, the switching was achieved by asymmetric thermal expansion of a U-shaped cantilever beam. Pure thermal expansion, mainly taking place in the silicon beam, which was heated by a deposited thin-film heater, caused the beam to deflect through the bimaterial effect occurring between the silicon and the heater material itself. An 8-channel switch was fabricated on a single silicon substrate using anisotropic wet etching in KOH. The switches exhibited an insertion loss of 1 dB and a -60 dB crosstalk.

Another interesting use of cantilever-beam MOEMs is in the field of sensors. For example, Ollier et al. [19] have made a vibration sensor in silicon that features an input waveguide that tracks along the surface of a micro-fabricated silicon cantilever beam before it is coupled by means of a multimode interference coupler into two output waveguides. At the end of the beam near the point of coupling to the output waveguides is a seismic mass of 2.5 μ g that moves under acceleration. The length of the beam is 600 μ m. All of the devices are integrated on a single silicon substrate. Both the input waveguide and the seismic mass are made of silica. The waveguide is a three-layer structure with different phosphorous doping levels to change the index of refraction as required. As the weighted end of the beam moves, the proportion of the lightwaves in the input waveguide that is coupled into each of the output guides changes. At zero displacement of the beam the output intensities are equal. The difference in intensity of the two output channels is proportional to acceleration. Accelerations in the range from 0.8 to 400 m/s² were measurable with a linearity better than 1% and a resolution of 0.5 m/s². The frequency range was measured to be 30–2000 Hz.

19.4 Torsional Devices

A torsional force is one that tends to cause rotation of a mechanical element around a particular axis. It is not necessary that the rotation be through a full 360° . In fact, most of the MOEM applications that employ torsional forces involve rotation through a smaller angle. One device that forms a convenient bridge between a discussion of cantilever-beam and torsional devices is the wavelength add-drop switch of Ford et al. [20]. An add-drop switch is a device used in wavelength division multiplexed telecommunication systems to add in a new information channel at some point along an optical fiber transmission line or to drop off a channel at some location to route it to the desired receiver. Whether adding or dropping a channel, the remaining channels must be undisturbed. These multiplexing switches are generally called optical add drop multiplexers (OADMs). An OADM has two input ports, “in” and “add”, and two output ports, “pass” and “drop”. In the particular

device under discussion, the light wave data stream enters the “in” port and passes through an optical circulator to a wavelength demultiplexer, and then on to an array of 16 tilt-mirror switches. The add-drop function is accomplished by means of an electrostatically actuated, tilting rectangular micromirror that tilts around an axis defined by two support points, one at each end. Depending on the angle of the mirror the lightwaves are reflected to either the “pass” or “drop” ports via a circulator. A signal entering the “add” port is similarly directed to the “pass” output port. This OADM uses free space optical wavelength multiplexing and a column of micromechanical tilt-mirrors to switch 16 channels, spaced at 200 GHz intervals between 1531 and 1556 nm. Each channel is switched by a single 1×1 tilt-mirror switch that either passes the lightwaves into the fiber transmission line (“pass” output port) or out of it (“drop” output port). The mirrors tilt to an angle of 5° with the substrate when an activating voltage of 20 V is applied. The fiber to fiber insertion loss to the “pass” output is 5 dB, and that to the “drop” output is 8 dB, with 0.2 dB polarization dependence. The switching speed is 20 μs . These MOEM chips were fabricated by the same “MUMP” foundry (operated by the Microelectronics Center of North Carolina) [9] that was used to fabricate the deformable mirror membrane [8] described in Section 19.2. The MUMP process is a general purpose polysilicon micromachining process.

In Section 19.3 a 2×2 optical switch [17] was discussed that used a cantilever beam to move a 45° angled mirror to switch an optical beam (see Fig. 19.6). Toshiyoshi et al. [21] have used the same configuration in an electromagnetically operated torsion mirror switch. In their switch, the mirror is supported by torsion beams at one edge that hold it in a horizontal position, clear of the optical beam. (The torsion beams are small segments of unremoved polysilicon.) The $150 \times 500 \mu\text{m}$ mirror of polysilicon is coated with a layer of the magnetic material FeNiCo, topped with a gold reflecting surface. When a magnetic field is applied normal to the rest position of the mirror by means of an electromagnet, the mirror rotates upward 90° to its reflecting vertical position. When the field is turned off, the mirror is returned to its horizontal rest position by the restoring force of the torsion beams. Switching time was measured to be in the range from 10–50 ms. Insertion loss was measured to be 2.5 dB in the reflecting state and 0.84 dB in the pass-through transmission state. Switching contrast of more than 45 dB and crosstalk less than -45 dB also were measured.

This type of magnetically actuated torsion plate can be used in a variety of different MOEMs. The following approach can be used to determine the magnetic field strength needed to produce a given amount of angular rotation of the torsion plate.

The magnetic torque T_m produced by an applied magnetic field strength H_d is given by [21]

$$T_m = V_m M_s H_d \sin\left(\frac{\pi}{2} - \theta\right), \quad (19.20)$$

where M_s is the saturation magnetization and V_m is the volume of magnetic material on the mirror.

The mechanical restoring torque of the torsion beams is

$$T_r = \frac{2(Gwt^3\theta)}{3l} \left(1 - \frac{192}{\pi^5}\right) \frac{t}{w} \tanh\left(\frac{\pi w}{2t}\right), \quad (19.21)$$

where l , w and t are the length, width and thickness of the beams and G is the sheet modulus of silicon (73 Gpa). The above equation assumes that $w \gg t$.

Setting (19.20) and (19.21) equal to each other gives the values of θ as a function of magnetic field strength H_d . In order to reduce power consumption, a latching mechanism has been proposed [21] that utilizes the natural hysteresis of the magnetization curve (i.e. residual magnetism) to hold the plate in the reflecting position.

A rotating, reflecting torsion plate also can be used to make an optical beam scanner. For example, Schenk et al. [22] have made an electrostatically-driven scanning mirror for scanning a laser beam. The device was fabricated on a silicon-on-insulator (SOI) wafer. The shape of the device is essentially that shown in Fig. 19.3(b), except that the torsion beams (springs) are electrically isolated from the support frame by etched trenches back-filled with high resistivity polysilicon. Back-and-forth rotational oscillation of the scanning mirror about the axis of the torsion springs is achieved by applying a suitably time varying voltage between the torsion plate and the supporting frame. The most effective driving waveform was found to be a rectangular pulse of 50% duty factor, with the frequency chosen so as to synchronize the turn-off of the pulse to occur when the mirror was passing the rest position. When the torsion plate is in the rest position (lying in the plane of the supporting frame) the electrostatic forces due to the applied voltage are also in that plane. Thus, an additional “starting” electrode is required on one side of the frame to create an asymmetric imbalance and start the oscillation. Once started, the torsion plate passes through this position because of its momentum and the oscillation continues. The oscillation of the mirror produces a one-dimensional (1-D) scanning of a laser beam directed onto its surface. However, 2-D scanning can be obtained by mounting the support frame of the 1-D scanning plate in a suitably driven gimbaled mount. This mount can be fabricated in monolithically integrated from with the basic device, as shown in Fig. 19.7, with electrical isolation provided so that both dimensions can be scanned independently. It was found that a 60° scan angle could be achieved with a 20 V driving voltage.

Resonance frequencies for these devices were measured to be approximately 1 kHz, and did not change after operation in run tests for up to 10^9 periods.

It is possible to use torsional forces to make a rotary electrostatic micromotor. A device of this type has been demonstrated by Azzam-Yasseen et al. [23] to be useful as a 1×8 optical switch, as shown in Fig. 19.8.

As the mirror turns, the optical beam is switched from one output fiber to the next. The device can be operated either as a stepping motor by energizing only the next pole, or as a continuously running motor by energizing the poles with three-phase excitation. The micromotor structure is made by fabricating both silicon and nickel parts using high-aspect-ratio microfabrication techniques and then assembling the parts. Single crystal silicon is used so that the motor can be integrated

Fig. 19.7 Two-dimensional scanner configuration

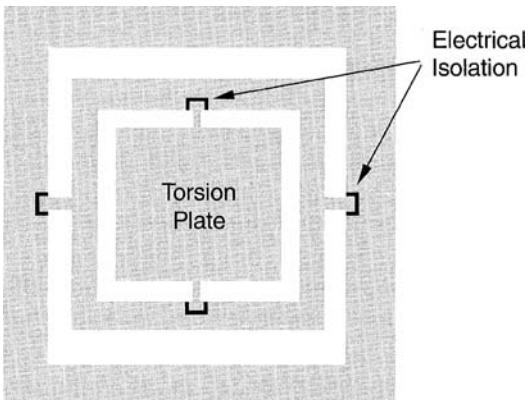
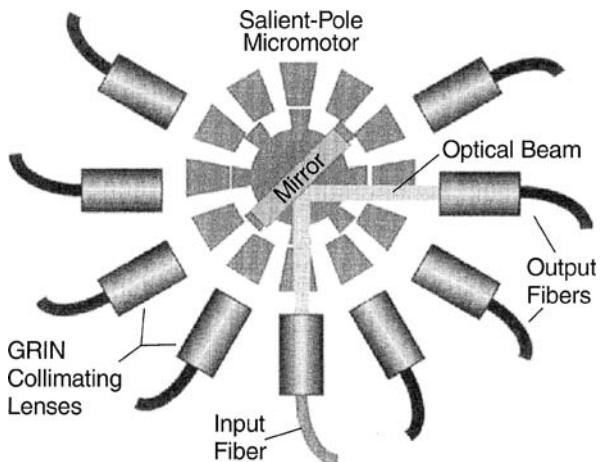


Fig. 19.8 Rotary micromotor switch [23] © 1999 IEEE



with electronic circuitry. The salient-pole micromotor has a 1 mm diameter rotor that is 200 μm thick. The rotor supports up to a 500 μm tall, 900 μm wide mirror while it rotates on a central bearing post. After microfabrication, individual parts were chosen carefully to match their dimensions. Combinations of silicon stators with nickel rotors were found to work best. The devices that functioned most satisfactorily had a rotor diameter of 995 μm , a central hole of 79 μm diameter, a stator diameter of 1029 μm , and a bearing post diameter of 68 μm . The assembly resulted in a rotor/stator gap of 17 μm with a 5 μm bearing clearance. When the device was operated as a stepping motor the switching time was observed to be on average 18 ms. When running continuously, the micromotor had a maximum speed of approximately 300 rpm. A typical operating voltage was 50 V. Optical switching was tested in both single and multi-mode fibers at a wavelength of 1310 nm, and in multimode fibers at 850 nm. Coupling loss as low as 0.96 dB was measured in the multimode case and 2.32 dB in single-mode. Crosstalk between channels was less than -45 dB.

19.5 Optical Elements

All of the MOEMs discussed so far have had moving parts. However, that is not always necessary. Many optical elements that can be made using the same micro-fabrication techniques have no moving parts. These include lenses, filters, couplers, mirrors, beam splitters, diffractive optical elements (DOEs) and holographic optical elements (HOEs). Microcylindrical lenses have been fabricated directly on the emitting surface of a 635 nm wavelength diode laser by depositing a carefully shaped layer of SiO_x [24]. The deposition was done by programmed focused-ion-beam (FIB) sputtering. Coupling efficiency as high as 81% into a single-mode, 10- μm core fiber was observed. The far field angle of the laser was reduced from a full-angle of 31° to only 2.1° when the lens was applied.

Very sophisticated lenses and diffractive elements can be fabricated in semiconductors and other device materials such as SiO_2 and nickel by means of focused ion beam sputtering. Since it is composed of charged particles, the ion beam can be focused by electrical and magnetic fields. Beam spot sizes on the order of 100 nm can be achieved. The position and tracking of the ion beam is controlled by a computer program. The optical elements can be either directly fabricated in the substrate material, or molds can be made by FIB milling in a hard material like nickel to permit replication of the optical element by filling or hot-embossing. Submicrometer features are achievable. Some good examples of what can be done by this method have been given by Yongqi and Bryan [25], as shown in Fig. 19.9.

The DOE of Fig. 19.9a clearly can be seen to have well defined submicrometer-sized features. The elements in (a) and (b) were both fabricated by direct FIB milling of the substrate, while the cylindrical microlens shown in (c) was made by FIB deposition of SiO_2 . The specific details of the FIB process depend on the features of the device that is being fabricated and the equipment being used. Generally a heavy ion that won't react chemically with the substrate is used, at an energy between 5 to 50 keV, and with a beam-limiting aperture of 25 to 400 μm . Lower ion beam energy does not produce adequate sputtering depth, and higher beam energy causes ion implantation rather than sputtering. The particular system used to produce the

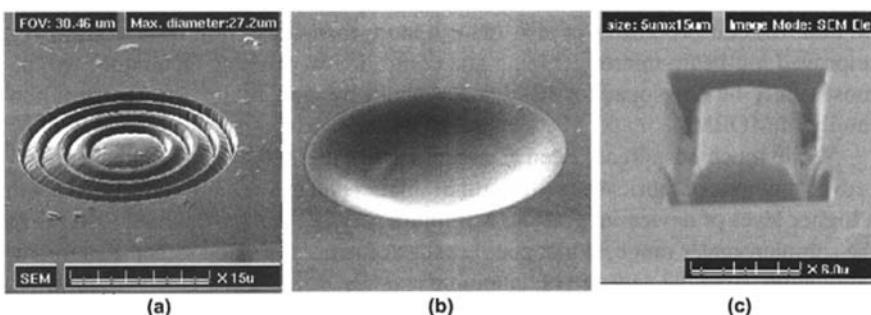
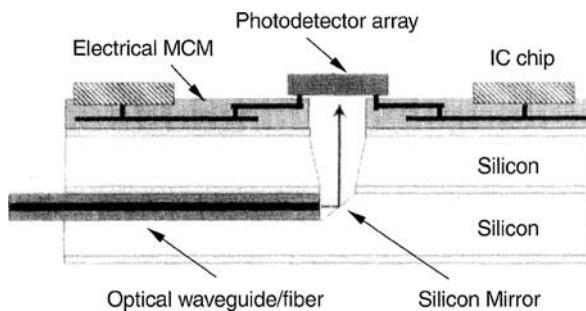


Fig. 19.9 a–c. Optical elements fabricated by FIB micromachining. **a** DOE with 8 μm diameter **b** Nickel microlens mold, dia. = 19 μm **c** 2.8 \times 7.5 μm cylindrical lens [25]

Fig. 19.10 I/O coupler with integrated optoelectronic components [26] ©1999 IEEE



devices in Fig. 19.7 employed a beam of 50 keV Ga^+ ions, with ion spot size of 215 nm. Operated in the raster scan mode it had a writing speed of $20 \mu\text{m}^2/\text{s}$.

The micromachining techniques of FIB sputtering and controlled selective wet etching can be combined with more conventional semiconductor fabrication methods to produce OEICs composed of different types of optical, mechanical and electronic devices. An example of this type of MOEM is the optical input/output (I/O) coupler made by Oh et al. [26], that combines electronic IC chips within a micro-circuit module (MCM) with a MSM photodiode array and a fiber-optic coupler, as shown in Fig. 19.10.

A key feature of the MOEM in Fig. 19.10 is the use of conductive polymer bumps formed on the contact pads of GaAs MSM detectors and side alignment pedestals electroplated with NiFe to facilitate flip-chip bonding of the MSM photodetector array. Alignment errors were estimated to be less than $\pm 5 \mu\text{m}$. In the silicon substrate, non-isotropic preferential etching was used to produce a (111)-oriented micromirror and a fiber-aligning V-groove to couple lightwaves from an input optical waveguide fiber. The conductive polymer flip-chip bonding method yields a very low contact resistance $\sim 10 \text{ m}\Omega$ even though it is a relatively low-temperature process ($\sim 170^\circ\text{C}$). Thus the optical characteristics of the photodetectors were not affected by mounting and remained good, with a dark current of about 10 nA and a responsivity of 0.33 A/W at 870 nm. It seems that the use of the structure shown in Fig. 19.9, should not be limited to just coupling light from a fiber into a photodetector. It is expected that the same approach, or a slight modification of it, would be generally useful in the application of coupling light from any edge-oriented waveguiding device to a surface-oriented device, or vice versa. For example, it could be used to couple the light emitted by a VCSEL to a rectangular waveguide in the substrate.

19.6 Future Directions in MOEMS Development

All of the MOEM devices that have been mentioned in this chapter have been fabricated in silicon, except for the hybrid GaAs/Si I/O coupler described in Section 19.5. This grows out of the fact that silicon, being an elemental semiconductor as opposed to a binary or ternary material, is easier to process without dissociation and

deterioration. Silicon processing is also more mature than that in the III–Vs. However, improved ion-beam micromachining and controlled wet etching techniques are being developed for III–V materials, leading to a new family of MOEMS. GaAs RF MEM switches have been reported [27].

While much has already been accomplished in the case of silicon, new and improved methods of fabrication will result in smaller and more elaborate MOEMS with a higher level of device integration. Minimum feature sizes have already shrunk to the submicrometer range, so that people speak of nanofabrication rather than micro-fabrication. A whole new area of “photonic crystals” has recently emerged, in which scientists and engineers are studying periodic structures with periodicity so fine that the existence of photonic transmission bands and forbidden bands analogous to those for electrons in a crystal have been observed [28–30]. Photonic crystals are described in detail in Chapter 22.

MOEMS and MEMs have not been confined to just the laboratory. They have also proven their worth in commercial applications. One of the earliest commercial MEMs actuators was display using a two-dimensional array of tilting micromirrors developed by Texas Instruments (TI) [31]. These components have demonstrated high yield and high reliability, so that they are now used in a commercially available 800×600 pixel projection display.

19.7 Mechanical Properties of Silicon

Most of the MEMs that have been made to date are based on silicon. Therefore it is worthwhile to list the numerical values of some of the key mechanical properties of that material as is done in Table 19.2. These will be useful for doing the problems at the end of this chapter.

Table 19.2 Mechanical properties of silicon

Young's Modulus (E)	=	190 GPa (1 GPa = 145 038 lb/in ² or 1.02×10^8 kg/m ²)
Sheer Modulus (G)	=	73 GPa
Poisson's Ratio (ν)	=	0.28
Density (ρ)	=	2.3 g/cm ³

Problems

- 19.1 If a silicon rod of 1 mm diameter is exposed to a tensional force by supporting a 3 g weight,
- What axial stress does the rod experience?
 - What is the axial strain?
 - What is the longitudinal strain?

- 19.2 A square membrane of silicon, 7 mm on a side and 150 μm thick, is exposed uniformly over its surface to a pressure of $1 \times 10^4 \text{ kg/m}^2$.
- What is the deflection at the center point of the membrane?
 - What is the maximum longitudinal stress?
 - What is the maximum transverse stress?
 - What is the resonant frequency of the membrane?
- 19.3 A silicon cantilever beam, 1.5 mm long, 500 μm wide and 250 μm thick, is loaded with a uniformly distributed force per unit width of 100 kg/m^2 .
- What is the deflection at the point 0.5 mm from the free end?
 - What is the maximum stress?
 - What is the frequency of the fundamental vibrational mode?
- 19.4 Repeat problem 19.3 if the beam is loaded with a point load of 10 g at the free end.
- 19.5 A silicon torsion plate is supported by (silicon) beams that are 200 μm long, 100 μm wide and 100 μm thick. What is the restoring torque that is developed when the plate is rotated through an angle of $\pi/4$ radians?
- 19.6 A round silicon membrane of diameter = 8 mm, and thickness = 200 μm is exposed to a uniform pressure of $1 \times 10^4 \text{ kg/m}^2$ over its surface. What is the deflection at the center of the membrane?
- 19.7 A silicon nitride cantilever beam is defined to be 1000 μm of length by the photomask. After the beam is released (separated from the substrate), the beam changes length due to the tensile stress. Calculate the length change if the tensile (axial) stress is -20 MPa (megaPascals). The Young's Modulus for silicon nitride is 280 GPa.
- 19.8 What are the three major categories of MOEM (MEM) devices? Give one example of each.
- 19.9 A round silicon membrane has a diameter of 12 mm and a thickness of 250 μm . If it is exposed to a uniform pressure of $1 \times 10^4 \text{ kg/m}^2$ over its surface, what is the deflection at a point located 4 mm from the center?

References

- E.P. Popov: *Introduction to the Mechanics of Solids* (Prentice Hall, Englewood Cliffs, NJ 1968)
- American Institute of Physics Handbook*, 3rd edn., D.E. Gray (ed.) (McGraw Hill, New York, 1972). K. Peterson: Silicon as a mechanical material. IEEE Proc. **70**, 420 (1982)
- G. Cibuzar: MEMs, in *The Science and Engineering of Microelectronic Fabrication*, S.A. Campbell (ed.), 2nd edn., (Oxford University Press, New York, 2001) Chap. 19
- H. Toshiyoshi, D. Miyauchi, H. Fujita: Electromagnetic torsion mirrors for self-aligned fiber-optic crossconnectors by silicon micromaching. IEEE J. Sel. Top. Quant. Electron. **5**, 10 (1999)
- W.G. Wu, D.C. Li, W. Sun, Y.L. Hao, G.Z. Yan, S.J. Jin: Fabrication and characterization of torsion-mirror actuators for optical networking applications, Sensors and Actuators A: Physical **108**, 175 (2002)

6. J. Zhoa, S. Dasgupta, H. Kobayashi, H.E. Jackson, J.T. Boyd: Optically interrogated MEMS pressure sensors for propulsion applications. *Opt. Eng.* **40**, 598 (2001)
7. H. Porte, V. Gorel, S. Kiryenko, J.-P. Goedgebuer, W. Daniau, P. Blind: Imbalanced Mach-Zehnder interferometer integrated in micromachined silicon substrate for pressure sensor. *IEEE J. Lightwave Tech.* **17**, 229 (1999)
8. T.G. Bifano, J. Perreault, R. Krishnamoorthy-Mali, M.N. Horenstein: Microelectromechanical deformable mirrors. *IEEE J. Sel. Top. Quant. Elect.* **5**, 83 (1999)
9. D. Koester, R. Mahadevan, K.W. Markus: MUMP's introduction and design rules. tech. paper, MCNC Technology Applications Center, 3021 Cornwallis Road, Research Triangle Park, NC, Oct, 1994
10. G.J. Veldhuis, T. Nauta, C. Gui, J.W. Berenschot, P.V. Lambeck: Electrostatically actuated mechanooptical waveguide ON-OFF switch showing high extinction at a low actuation voltage. *IEEE J. Sel. Top. Quant. Elect.* **5**, 60 (1999)
11. F. Chollet, M. de Labachelerie, H. Fujita: Compact evanescent optical switch and attenuator with electromechanical actuation. *IEEE J. Sel. Top. Quant. Elect.* **5**, 52 (1999)
12. C. Marxer, M.-A. Gretillat, N.F. de Rooij, R. Battig, O. Anthamatten, B. Valk, P. Vogel: Reflective duplexer based on silicon micromechanics for fiber-optic communication. *IEEE J. Lightwave Tech.* **17**, 115 (1999)
13. E.C. Vail, M.S. Wu, G. Li, W. Yuen, C.J. Chang-Hasnain: Micromachined wavelength tunable optoelectronic devices with record tuning. *Electronics Lett.* **19**, 1671 (1995)
14. E.C. Vail, M.S. Wu, G. Li, W. Yuen, C.J. Chang-Hasnain: Widely and continually tunable resonant cavity detector with wavelength tracking. *IEEE Photonics Tech. Lett.* **8**, 98 (1996)
15. E.C. Vail, M.S. Wu, G. Li, W. Yuen, C.J. Chang-Hasnain: GaAs micromachined widely tunable Fabry-Perot filters. *Electron. Lett.* **31**, 228 (1995)
16. C.J. Chang-Hasnain: Tunable VCSEL. *IEEE J. Sel. Top. Quant. Elect.* **6**, 978 (2000)
17. R.T. Chen, H. Nguyen, M.C. Wu: A high-speed low-voltage stress-induced micromachined 2 × 2 optical switch. *IEEE Photonics Tech. Lett.* **11**, 1396 (1999)
18. M. Hoffmann, P. Kopka, E. Voges: All-silicon bistable micromechanical fiber switch based on advanced bulk micromachining. *IEEE J. Sel. Top. Quant. Elect.* **5**, 46 (1999)
19. E. Ollier, P. Philippe, C. Chabrol, P. Mottier: Micro-opto-mechanical vibration sensor integrated on silicon. *IEEE J. Lightwave Tech.* **17**, 26 (1999)
20. J.E. Ford, V.A. Aksyuk, D.J. Bishop, J.A. Walker: Wavelength add-drop switching using tilting micromirrors. *IEEE J. Lightwave Tech.* **17**, 904 (1999)
21. H. Toshiyoshi, D. Miyauchi, H. Fujita: Electromagnetic torsion mirrors for self-aligned fiber-optic crossconnectors by silicon micromachining. *IEEE J. Sel. Top. Quant. Elect.* **5**, 10 (1999)
22. H. Schenk, P. Durr, T. Haase, D. Kunze, U. Sobe, H. Lakner, H. Kuck: Large deflection micromechanical scanning mirrors for linear scans and pattern generations. *IEEE J. Sel. Top. Quant. Elect.* **6**, 715 (2000)
23. A. Azzam-Yasseen, J.N. Mitchell, J.F. Klemic, D.A. Smith, M. Mehregany: A rotary electrostatic micromotor 1 × 8 optical switch. *IEEE J. Sel. Top. Quant. Elect.* **5**, 26 (1999)
24. F. Yongqi, N.K.A. Bryan, O.N. Shing: Integrated micro-cylindrical lens with laser diode for single-mode fiber coupling. *IEEE Photonics Tech. Lett.* **12**, 1213 (2000)
25. F. Yongqi, N.K.A. Bryan: Investigation of direct milling of micro-optical elements with continuous relief on a substrate by focused ion beam technology. *Opt. Eng.* **39**, 3008 (2000)
26. K.W. Oh, C.H. Ahn, K.P. Roenker: Flip-chip packaging using micromachined conductive polymer bumps and alignment pedestals for MOEMS. *IEEE J. Sel. Top. Quant. Elect.* **5**, 119 (1999)
27. S.C. Shen, D. Caruth, M. Feng: Broadband low actuation voltage RF MEM switches, Digest IEEE 2000 GaAs IC Symposium, Seattle, WA, (Nov. 5–8, 2000)
28. E. Yablonovitch: Inhibited spontaneous emission in solid-state physics and electronics. *Phys. Rev. Lett.* **58**, 2059 (1987)
29. J.D. Joannopoulos, R.D. Meade, J.N. Winn: *Photonic Crystals, Molding the Flow of Light*. (Princeton University Press, Princeton, NJ 1995)

30. H. Benisty, C. Weisbuch, D. Labilloy, M. Rattier, C.J.M. Smith, T.F. Krauss, R.M. de la Rue, R. Houdre, U. Oesterle, C. Jouanin, D. Cassagne: Optical and confinement properties of two-dimensional photonic crystals. *IEEE J. Lightwave Technol.* **17**, 2063 (1999)
31. P.F. Van Kessel, L.J. Hornbeck, R.F. Meier, M.R. Douglass: A MEMs-based projection display. *Proc. IEEE* **86**, 1687 (1998)

Chapter 20

Applications of Integrated Optics and Current Trends

In the preceding chapters, the theory and technology of optical integrated circuits have been described. Although this a relatively new field of endeavor, numerous applications of OIC's to the solution of current engineering problems have already been implemented and some OIC's are now available as "off-the-shelf" commercial products. Of course, optical fiber waveguides, the companion element of OIC's in an integrated-optic system, are already well recognized as being very useful consumer products. In this chapter, some of the more recent applications of both fibers and OIC's are reviewed, and current trends are evaluated. In this review of representative integrated optic applications, specific systems and companies are named in order to illustrate the international character of the field and the types of organizations that are involved in it. Recommendation of any particular company or its products is not intended or implied. Also, the performance data that are quoted have generally been obtained from news articles and other secondary sources. Hence, they should be interpreted as being illustrative rather than definite.

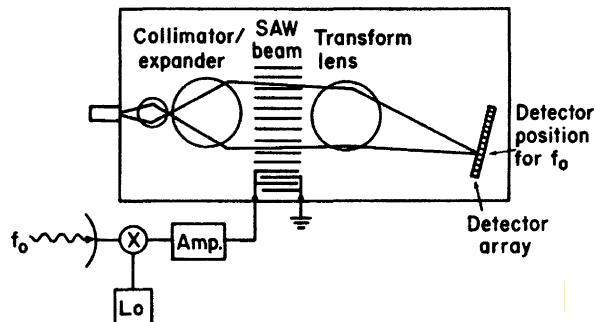
20.1 Applications of Optical Integrated Circuits

20.1.1 RF Spectrum Analyzer

Probably the earliest demonstration of a multi-element OIC that was performed was the hybrid implementation of the real-time rf spectrum analyzer, which was originally proposed by Hamilton et al. [1]. The purpose of this spectrum analyzer, is to enable the pilot of a military aircraft to obtain an instantaneous spectral analysis of an incoming radar beam, in order to determine if his plane is being tracked by a ground station, air-to-air missile, etc. Obviously, such information is required if he is to be able to quickly take effective evasive action. Of course, the frequency content, or *signature*, of all enemy radar signals that are likely to be encountered would have to be available for comparison, probably stored in the memory of the plane's onboard computer.

A diagram of the integrated-optic spectrum analyzer is shown in Fig. 20.1. Light from a laser source is coupled into a planar waveguide, in which it passes first

Fig. 20.1 Diagram of an integrated-optic rf spectrum analyzer



through a collimating lens, then through a Bragg-type acousto-optic modulator. The rf signal to be spectrally analyzed is applied to the acoustic transducer that generates the sound waves, causing them to have a time varying period. Thus the deflection angle of the optical beam at the output of the modulator is a function of the rf signal. A second lens is used to focus the optical beam onto an array of photodetectors. If more than one frequency component is present in the rf signal, the light beam is divided into corresponding components that are focused onto different detector elements. Each detector element represents a particular frequency channel, and, since photo-diodes generally have square law response, the output signal from any channel is proportional to the rf power at that frequency. The advantage of an integrated-optic spectrum analyzer, as compared to an electronic one, is that only a few optical elements are needed to perform a function that would otherwise require thousands of electronic elements.

The development of working models of the integrated optic rf spectrum analyzer took place at several different laboratories and extended over a number of years. The first working model was produced by the Westinghouse Advanced Technology Laboratories in 1980 [2, 3]. It was fabricated on an X-cut LiNbO₃ substrate, approximately $7 \times 2.5 \times 0.3$ cm³, in which a planar waveguide had been produced by indiffusion of titanium at 1000°C. The lenses used were of the geodesic variety, formed by machining *dimples* into the surface of the substrate prior to waveguide diffusion. In this type of lens, light waves are still confined by the waveguide, but they follow the longer curved path through the lens region. Since waves traveling near the center of the lens go over a greater path length than waves traveling near its edges, the wavefront is modified, so that focusing can occur. Such lenses can be made with surprising accuracy. The two aspherically-corrected geodesic lenses in the Westinghouse spectrum analyzer had essentially diffraction-limited spot sizes. The silicon diode detector array contained 140 elements, and was butt coupled to the waveguide. Design parameters for the Westinghouse integrated-optic spectrum analyzer are shown in Table 20.1. The input lens focal length was chosen to expand a 6 μm GaAlAs laser spot to 2 mm by diffraction.

The spectrum analyzer was tested first with a He-Ne laser source of 6328 Å wavelength and found to have a bandwidth of 400 MHz with a resolution of 5.3 MHz.

Table 20.1 Westinghouse spectrum analyzer, design parameters [2, 3]

Substrate size	$7.0 \times 2.5 \text{ cm}^2$
Front face to collimating lens	2.45 cm
Collimating lens diameter	0.80 cm
Collimating lens focal length	2.45 cm
Spacing between lenses	1.80 cm
Transform lens offset angle	3.79°
Transform lens offset	0.06 cm
Transform lens focal length	2.72 cm
Detector array pitch	12 μm
Laser beam width	6 μm
Number of detector elements	140
SAW transducer type	2 element, tilted

Later results obtained by using a butt-coupled GaA1As laser diode emitting at 8300 Å as the source showed an improved resolution of 4 MHz. Other performance characteristics are given in Table 20.2. The 400 MHz bandwidth limitation is mostly caused by the acoustic transducer, and may be improved by using a more sophisticated transducer, as described in Chapter 10. In any case, the spectrum analyzer could be used over a wider frequency range by using a local oscillator and mixer at the input to the transducer, as shown in Fig. 20.1. Thus, heterodyning could be used to electronically shift the 400 MHz bandpass to various center frequencies as desired.

Shortly after a working model of an rf spectrum analyzer was demonstrated by Westinghouse, an alternate embodiment of essentially the same design was demonstrated by Hughes Aircraft Company. The Hughes OIC also followed the basic pattern proposed by Hamilton et al. [1], as shown in Fig. 20.1. However, it differed from the initial Westinghouse OIC in that it featured a butt-coupled GaA1As laser diode rather than a He-Ne laser source, and the detector array was composed of silicon charge-coupled-devices (CCD) [4, 5] rather than photodiodes. The Hughes spectrum

Table 20.2 Westinghouse spectrum analyzer, performance characteristics [2, 3]

Center frequency	600 MHz
Frequency bandwidth	400 MHz
Frequency resolution	
with He-Ne 6328Å source	5.3 MHz
with GaA1As 8300Å source	4 MHz
Detector integration time	2 μs
Detector element spacing	12 μm (with no dead space between elements)
Full width at half power of focused spot in detector focal plane	3.4 μm ($1.02 \times$ diffraction limited size)
Bragg diffraction efficiency	50 to 100%/w

analyzer exhibited a 3 dB bandwidth of 380 MHz with a diffraction efficiency of 5% (at 500 mW rf power) [6, 7]. Operating at a wavelength of 8200 Å, the OIC had a resolution of 8 MHz and a linear dynamic range greater than 25 dB. Losses in the two geodesic lenses were measured to be less than 2 dB each.

The two embodiments of the rf spectrum analyzer described above are excellent examples of hybrid optical integrated circuit technology. By fabricating the laser diode in GaAlAs, the detector array in silicon, and the Bragg modulator in LiNbO₃, one can use the best features of all three materials to advantage. The major disadvantage of the hybrid approach is that all of these substrate materials must be carefully aligned and permanently bonded with micrometer-tolerance precision. Thermal expansion and vibration must somehow be prevented from destroying the alignment. Despite these difficulties, hybrid OIC's have been demonstrated to be viable structures, and will continue to be used in many applications even after monolithic technology has been fully developed.

20.1.2 Monolithic Wavelength-Multiplexed Optical Source

One of the applications for which optical integrated circuits were proposed early in the history of the field is an optical-frequency-multiplexed transmitter, such as that shown previously in Fig. 1.1, in which a number of DFB lasers, operating at different wavelengths, are coupled into a single fiber transmission line. An OIC of this type has, in fact, been fabricated by Alki et al. [8], using GaAlAs monolithic technology. Six DFB lasers, operating at wavelengths separated by 20 Å, were fabricated on a 5 mm square GaAs substrate by a two-step LPE growth process. The lasers had a separate confinement heterostructure (SCH) [9]. Third-order gratings were made on the surface by chemical etching, by using a mask made by holographic lithography. The lasers were coupled to undoped Ga_{0.9}Al_{0.1}As waveguides by direct transmission, as shown in Fig. 20.2. The lateral dimensions of the lasers and waveguides were defined by mesa etching down to the GaAs substrate to produce stripes that were 20 μm wide and 3 μm thick. The separation of the lasers was 300 μm, and the waveguides were curved through bends of minimum radius equal to 4 mm, in order to bring them together in a confluent coupler, as shown in Fig. 20.3. The output of the coupler was obtained via a single waveguide that was butt coupled to an optical fiber.

The lasers were operated by applying 100 ns current pulses at a repetition rate of 1 kHz. The differential quantum efficiency of the lasers was measured to be 7%, and the waveguide loss coefficient was about 5 cm⁻¹. The threshold current densities of the lasers were in the range from 3 to 6 kA/cm² at room temperature. The wavelength separation between lasers was measured to be 20 ± 5 Å. No difficulty was encountered in separately modulating the six lasers, and the overall differential quantum efficiency, measured at the launching output terminal, was about 30%. Thus, this early version of a monolithic chip represented a usable OIC, even though further refinements were to yield better efficiency.

Work on monolithic wavelength-multiplexed optical sources employing DFB lasers has continued over the years, with more recent work being directed toward

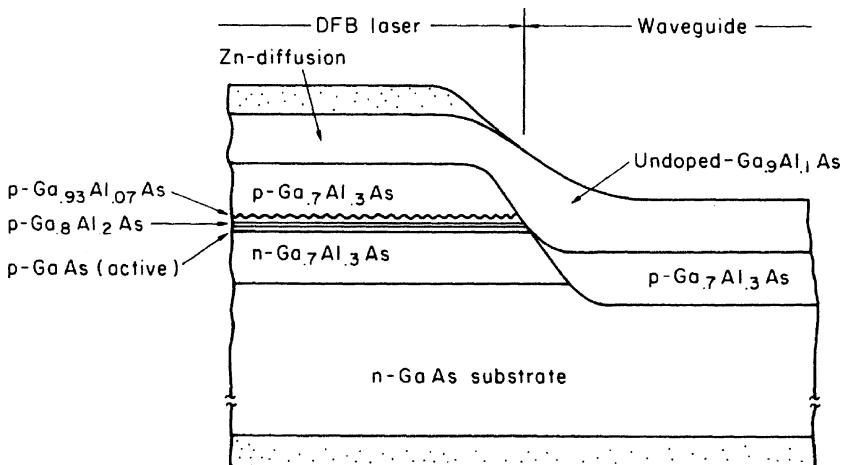


Fig. 20.2 DFB laser coupled to a GaAlAs waveguide by direct transmission [9]

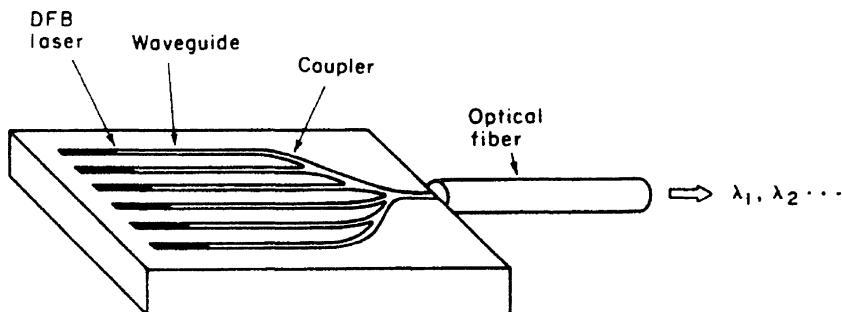


Fig. 20.3 Schematic drawing of a wavelength-multiplexed light source [8]

1.3 μm and 1.55 μm wavelength lasers for optical-fiber telecommunication systems. Zah et al. [10] have reported a study of wavelength-division multiplexed lighwave systems which has led to the conclusion that, in order to be cost effective, it is necessary to fabricate multi-wavelength laser transmitters by monolithic integration on one chip to reduce the cost of packaging and control circuitry by sharing them among all of the wavelengths. An example of such monolithic integration is provided by the frequency-division multiplexed ten-channel tunable DFB laser array of Sato et al. [11]. The lasers are tunable, multi-section, quarter-wave-shifted, strained InGaAsP MQW devices. The lasing frequencies of channels are spaced within a 10 GHz range. The linewidth of each channel is less than 2.3 MHz. A monolithically integrated chip containing 21 DBR lasers has been produced by Lee et al. [12]. They used sampled grating distributed Bragg Reflectors, in which the grating is not continuous, but rather is formed in bursts so that it has two inherent periodicities. This type of DBR grating makes the selection of laser emission wavelength more

accurate. The 21 InP/InGaAsP MQW lasers on the chip had emission wavelengths spanning over 40 nm with 0.8 nm spacing, centered at approximately 1.56 μm .

20.1.3 Analog-to-Digital Converter (ADC)

An analog-to-digital conversion method, proposed by Taylor [13, 14], has been implemented by Yamada et al. [15] in an optical integrated circuit that is capable of one-bit electro-optical AD conversion at a 100 MHz rate. The OIC incorporates two 3-dB couplers and a phase shifter, formed in a pair of straight waveguides, as shown in Fig. 20.4. The waveguides were fabricated by Ti diffusion of a LiNbO_3 substrate. The phase shifter was formed by a Ti double-diffusion, as shown in Fig. 20.4. An Al_2O_3 1100 Å thick buffer layer was used, separated by a 4 μm , gap between the waveguides to suppress dc drift. Waveguide spacing was 5.4 μm and device length was about 2 cm.

The configuration of two electro-optic couplers and a phase shifter forms a balanced-bridge modulator, with two complementary outputs which are equally affected by fluctuation of the light source. Hence a serious source of conversion error is inherently eliminated in this OIC. The integrated ADC was operated with a 1.15 μm -wavelength He-Ne laser source at bit rates up to 100 MHz. This initial success of high speed analog-to-digital conversion points the way towards more sophisticated, multi-bit, and monolithic OIC's. However, much remains to be done, especially in regard to development of a monolithic, high-speed, electronic or optical comparator to be incorporated into a fully monolithic ADC system.

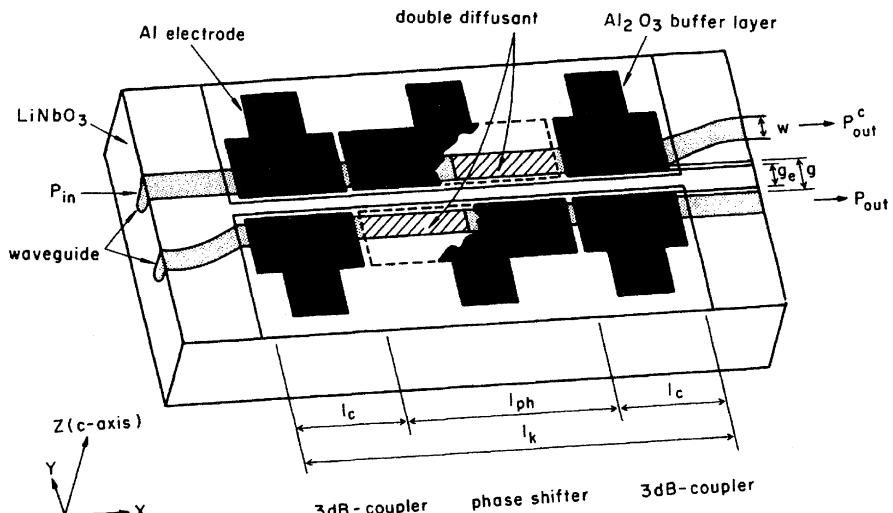


Fig. 20.4 Diagram of an integrated-optic analog to digital converter [15]

As work continues on these devices, optoelectronic ADCs capable of multibit conversion have been developed [16–21]. A twelve bit ADC developed by Twichell and Helkey [21] relies on a gain-switched diode laser and a dual-output Mach-Zehnder interferometer to produce phase-encoded sampling. Currie et al. [19] have produced a photonic ADC that maps an analog input waveform into a binary output by employing phase modulators and polarization-based optics.

20.1.4 Integrated-Optic Doppler Velocimeter

An integrated-optic Doppler velocimeter which employs both an optical fiber link and an OIC to measure velocity has been demonstrated by Toda et al. [22], as shown in Fig. 20.5. The optical integrated circuit was fabricated in a z-propagation LiNbO_3 substrate with Ti diffused waveguides. Laser beam lithography, with $0.2 \mu\text{m}$ accuracy, was used for waveguide patterning. The light source was a linearly polarized He-Ne laser. TE polarized light was focused into the input waveguide by a 20% lens and then split by a Y-branch coupler into a signal beam and a reference beam. In order to keep the outgoing beams separate from the reflected return beams. TE polarization was maintained on outgoing light while TM polarization was used for reflected light. Of course, polarization-maintaining optical fiber was used. In the case of the reference beam the TE/TM conversion was accomplished by means of an electro-optic mode converter. For the signal beam a quarterwave plate was used to produce TE/TM mode conversion. An absorptive TE/TM mode splitter was used to route the return signal beam to the avalanche photodiode (APD) which served as

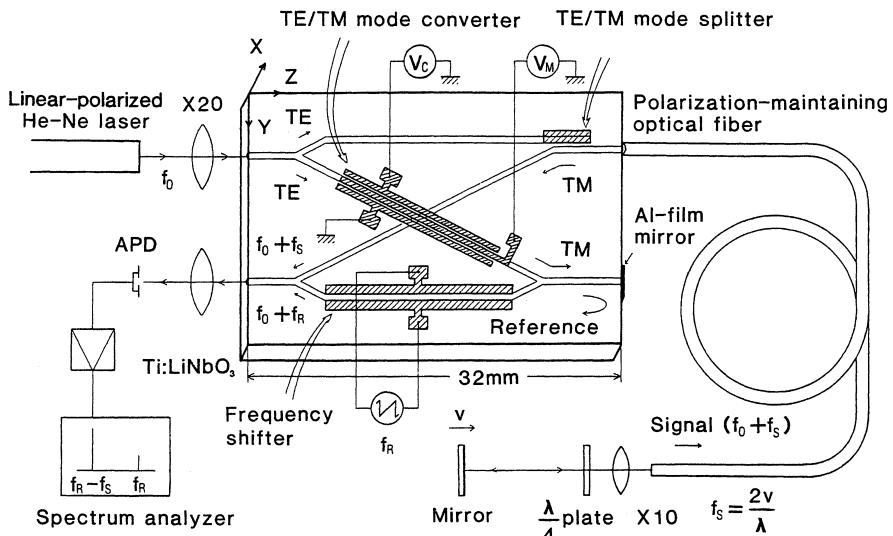


Fig. 20.5 An integrated-optic Doppler velocimeter

a mixer and detector. An electro-optic modulator was used to impress the reference modulation frequency f_R onto the reference beam. The Doppler effect produced a shift in the signal beam frequency from f_0 to $f_0 + f_s$, where f_s is given by

$$f_s = \frac{2v}{\lambda_0}, \quad (20.1)$$

where v is the velocity and λ_0 is the vacuum wavelength. After being recombined by a Y-branch coupler the reflected signal beam and reference beam were mixed in the avalanche diode. Because of the nonlinear response characteristic of the APD (it is a square-law device, i.e., response is proportional to the square of the electric field strength) the output photocurrent contained a beat frequency component $f_R - f_s$. Thus f_s , and hence the velocity, was determined. For a velocity of 8 mm/s the measured f_s was 25 kHz, and the signal to noise ratio was 25 dB.

This Doppler velocimeter demonstrates that integrated-optic technology can provide the compact and rugged heterodyne optics need for highly accurate measurement of velocity and displacement. By adding a balanced-bridge waveguide optical switch to an interferometric circuit of this type Toda et al. [23] have produced a time-division-multiplexed Doppler velocimeter which can measure two-dimensional velocity components v_x and v_y .

20.1.5 An IO Optical Disk Readhead

Optical disk information storage has found widespread use for computer data, as well as for video and audio reproductions. High data density and low background noise are key advantages of this method. However, relatively sophisticated optics must be used to insure good resolution and tracking of the light beam that is used to read information off the disk. For example, the optical readheads used in commercially available audio compact disk (CD) players often have eight or nine discrete optical elements, all of which have to be held in exact alignment in the face of much shock and vibration.

As an alternative, an integrated-optic optical disk pickup device capable of detecting readout and focus/tracking error signals has been designed and fabricated by Ura et al. [24], as shown in Fig. 20.6. The OIC was formed by depositing a planar # 7059 glass waveguide on a SiO_2 buffer layer on a silicon substrate. The light source was a butt-coupled GaAlAs laser diode. A chirped and curved focusing grating pattern coupler fabricated by electron beam direct writing lithography was used to focus the beam onto the disk, as well as to refocus the reflected beam back into the waveguide. A twin-grating focusing beam splitter served to divide the reflected beam into two beams which were focused onto two pairs of photodiodes formed in the Si substrate.

In operation, the pickup head provides not only a readout signal but also focus and tracking error signals. When the light beam is focused, the return beams hit both diodes of each pair equally. If the readhead is too close to the disk, the beams

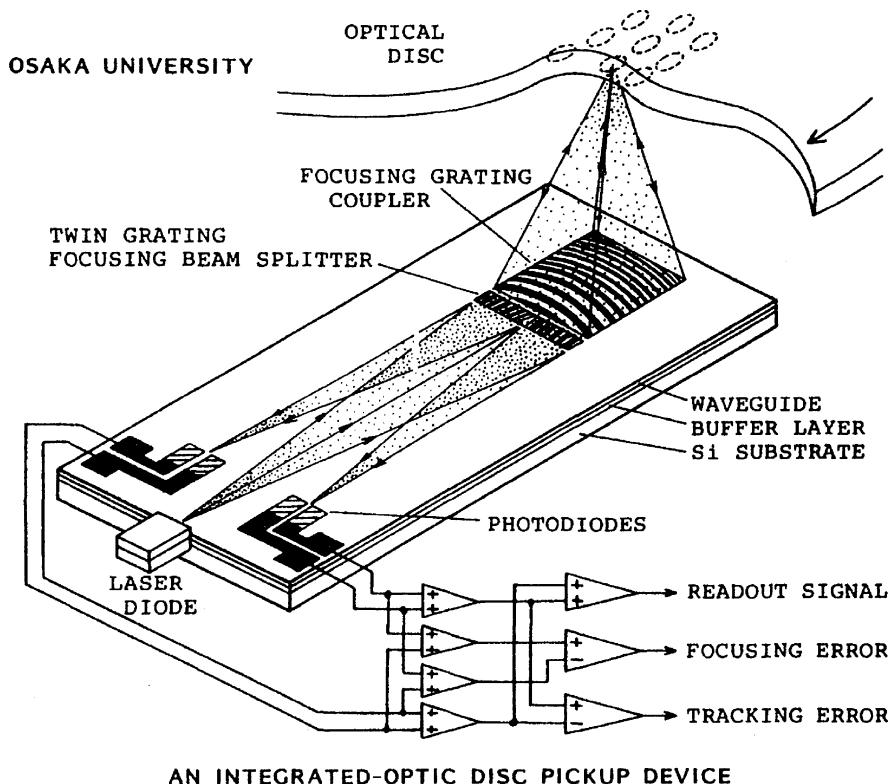


Fig. 20.6 An integrated-optic disc pickup device [24]

fall more on the outer diodes, while if it is too far away they fall more on the inner diodes. Tracking error is detected when the total intensity of the return beam reaching the left pair of diodes is not equal to that reaching the right pair. Thus conventional electronic comparators, sensing the photo-currents from the diodes, can be used to develop error signals to drive position correcting actuators.

The OIC readhead of Fig. 20.6, which has dimensions of only 5×12 mm, obviously has the advantage of being relatively insensitive to shock and vibration, as compared to a readhead fabricated from discrete optical components. While this OIC was first proposed as an optical disk pickup device, the same basic arrangement can be used more generally as a fully integrated interferometer position/displacement sensor with direction discrimination [25]. Such an interferometric sensor would be useful in a variety of high-precision positioning applications in which submicrometer accuracy is required.

An integrated optical disk readout head has also been made by Hudgings et al. [26]. Their device makes use of a vertical cavity surface emitting laser with an intra-cavity quantum well absorber. Detection of the reflected optical signal is performed by measuring the change in absorber voltage as the optical feedback into the VCSEL

cavity varies. The head has a $0.22 V_{pp}$ response Its RC time constant is $20 \mu s$, indicating a 50 kHz rolloff frequency.

Manoh et al. [27] have reported an integrated optical head device using a blue-violet laser diode. This head, which integrates seven optical elements and semiconductor chips into a $11 \text{ mm} \times 6 \text{ mm} \times 4.1 \text{ mm}$ package, is a key device for realizing a small and thin Blu-ray Disc drive.

20.1.6 OIC Temperature Sensor

The integrated-optic temperature sensor [28] shown in Fig. 20.7, requires no electrical connection, making it particularly useful in explosive or flammable environments in which an electrical sensor might be dangerous. The OIC is fabricated in a LiNbO_3 substrate with Ti diffused waveguides. It features a parallel array of three unequal arm-length Mach-Zehnder interferometers. The optical transmission of each interferometer varies sinusoidally with temperature, as shown in Fig. 20.7b, with a period which is inversely proportional to the optical path length difference between the two arms.

The optical transmission $P_{\text{out}}/P_{\text{in}}$ at the wavelength λ depends on both the effective index n_{eff} and the path length difference ΔL and is given by [28]

$$\frac{P_{\text{out}}}{P_{\text{in}}} = \frac{\gamma}{z} \left[1 + m \cos \left(\frac{2\pi}{\lambda} b \Delta L T + \Delta\phi_0 \right) \right], \quad (20.2)$$

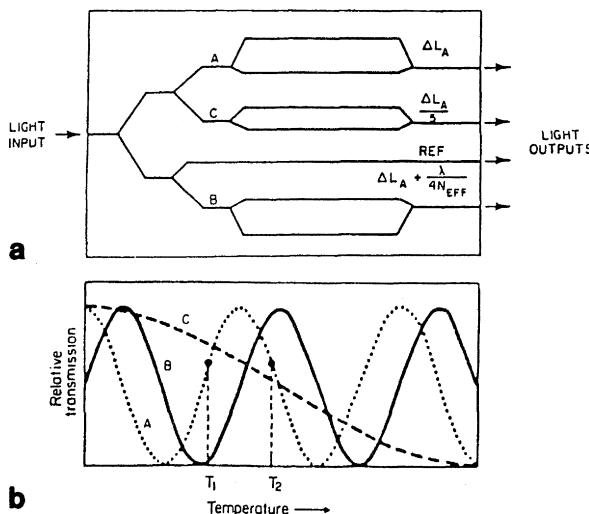


Fig. 20.7 a,b OIC temperature sensor **a** device structure **b** optical transmission characteristics [28]

where the constant of proportionality b is given by

$$b = \frac{dn_{\text{eff}}}{dT} + \frac{n_{\text{eff}}}{\Delta L} \frac{d(\Delta L)}{dT}. \quad (20.3)$$

Both n_{eff} and ΔL are functions of temperature T . The quantities γ and m are related to the insertion loss and depth of modulation of the interferometer, respectively. (For an ideal device $\gamma = m = 1.0$.) $\Delta\phi_0$ is a constant for a given device.

By measuring the transmission of all three interferometers one can determine the temperature. Two of the interferometers (A and B) have arm-length differences that are almost the same. Thus their transmission curves track close to one another, providing a high resolution in the temperature measurement. The third interferometer (C) has an arm-length difference which is only approximately one-fifth those of A and B. Thus one can determine which peak of the A and B transmission curves is being measured and temperature measurements can be made over a wide range. It has been reported that this temperature sensing OIC can measure with an accuracy of $2 \times 10^{-3}^\circ\text{C}$ over a 700°C range, when used with a 6328 \AA He-Ne laser as the light source. The sensor, of course, would be mounted at the point at which temperature measurement was desired and the optical input and output would be via optical fiber. Since the OIC chip is about 1 cm on a side, temperature measurements can be made on relatively small objects. Since the measurement signal is entirely optical, this device is relatively immune to electrical noise.

20.1.7 IO High Voltage Sensor

Integrated-optic Mach-Zehnder interferometers can be used to sense high voltage as well as temperature. A diagram of such a device is shown in Fig. 20.8 [29]. The waveguides are formed by Ti diffusion into a LiNbO_3 substrate. In this circuit the

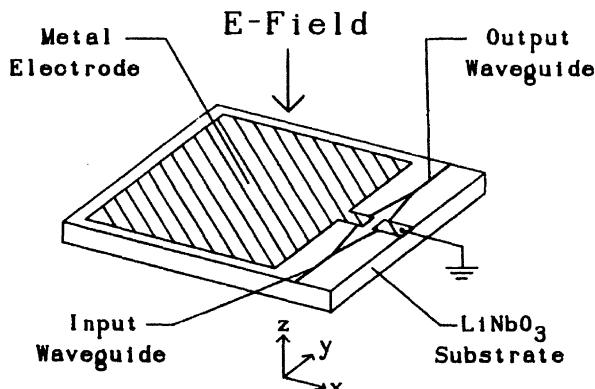


Fig. 20.8 IO high voltage sensor [29]

two branches of the interferometer are covered by metal electrodes which form a capacitive voltage divider. The electric field generated by the high voltage source induces a voltage on these electrodes which causes a relative phase shift between the optical waves in each arm, resulting in an intensity modulation of the output beam. The voltage-in/optical power-out transfer function is given by [29]

$$P_{\text{out}} = \frac{\alpha P_{\text{in}}}{2} \left[1 + \gamma \cos \left(\frac{\pi V}{V_\pi} + \phi_i \right) \right], \quad (20.4)$$

where P_{in} is the input power, ϕ_i is the intrinsic or zero voltage phase difference. V is the applied voltage and V_π is the half-wave phase shift voltage. The constants α and γ must be determined for a particular device. (For a perfect device $\alpha = \gamma = 1.0$.) Once the calibration curve has been determined for a particular sensor, the voltage can be accurately measured.

Since this sensor operates on the principle of an induced voltage it is not necessary to make electrical contact to the high voltage source, and input to, and output from the sensor can be via optical fiber. Thus good high voltage isolation can be maintained. The immunity to electrical noise provided by an optical fiber link is also a particularly important advantage when operating in a high voltage environment. This integrated optical high voltage sensor could be used, for example, for monitoring line voltages in SF₆-gas insulated bus ducts such as are used in power plants and switching stations.

20.1.8 IO Wavelength Meters and Spectrum Analyzers

By carefully arranging wavelength-selective optical elements in an OIC it is possible to make instruments that can measure an emission wavelength or spectrum. For example, Nabiev et al. [30] made a spectrophotometer that consists of two p-n junction InGaAs/GaAs QW photodiodes stacked vertically with a DBR reflector between them, as shown in Fig. 20.9a. The DBR reflector is composed of 10½ pairs of AlAs/GaAs. Lightwaves entering the top surface pass through both photodiodes and are detected by them. However, because of the wavelength selective reflectance of the DBR, some of the incident light doesn't reach the bottom photodiode. The ratio of the top detector response to that of the bottom diode is given by:

$$\frac{I_{\text{top}}}{I_{\text{bot}}} = R_0 \frac{1 + R(\lambda)}{1 - R(\lambda)}, \quad (20.5)$$

where R_0 is a constant of proportionality. The ratio of (20.5) is a single-valued function of wavelength that can be used to measure it, as shown in Fig. 20.9b, where the dotted curves are I_{top} and the solid curves are I_{bot} . The curve in (c) is a superposition of measurements at the three optical power levels shown in (b), indicating that measured wavelength is not sensitive to optical power.

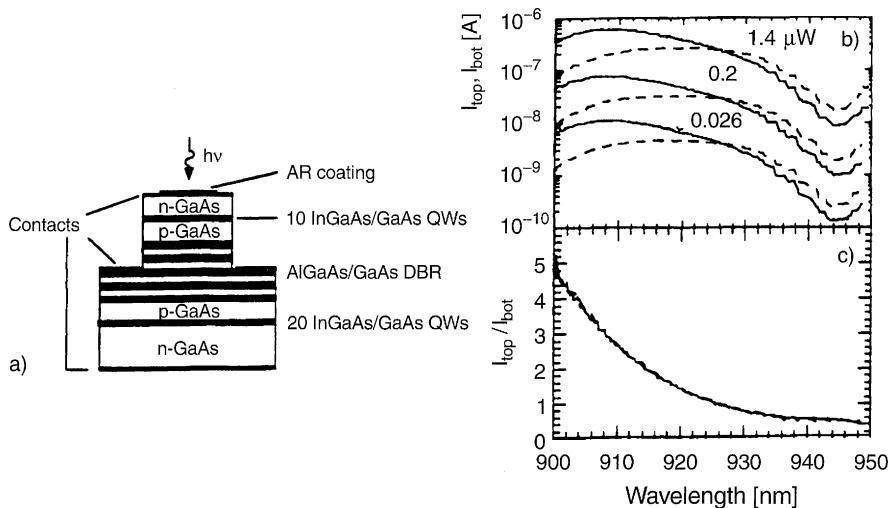


Fig. 20.9 a,b,c. Wavelength meter **a** device structure **b** response curves **c** response ratio [30]
© 1995 IEEE

An integrated optic device also can be used to measure an optical spectrum as well as just a single wavelength. For example, an OIC optical spectrum analyzer has been described by Madsen et al. [31]. Their device uses a tilted, chirped grating in a planar single-mode waveguide to diffract lightwaves of different wavelengths through different angles as they are coupled out of the waveguide. The light radiated out of the waveguide is captured by a vertical slab waveguide that directs it to a linear array of photodetectors. Each detector intercepts light of only a narrow range of wavelengths, so the optical spectrum of the light can be determined. The full-width half-maximum resolution was 0.15 nm over a bandwidth range of 7.8 nm. In this device the grating, which has a chirp of -1.75 nm/cm and a focal length of 12 cm, performs both the spatial dispersion and imaging (focusing) functions.

20.1.9 IO Chemical Sensors

Integrated optic devices can be used to sense the presence and the concentration of various chemical elements. These sensors generally function by measuring the change in some optical property of the material in a waveguide that is produced by the presence of the chemical to be sensed. Both absorption spectro-photometry and attenuated total reflection spectrometry can be used, as well as fluorescence spectrometry. The waveguide may be either an optical fiber or a multilayer guide in an OIC. An example of this type of sensor is the thin-film polyvinyl chloride (PVC) co-polymer sensor reported by Kim et al. [32]. When doped with (HDOPP-Ca), a neutral ionophore for calcium, dioctyl phthalate (DOP), and the chromoionophore (ETH5294) the PVC waveguide showed enough chemical reaction to the Ca^{2+} ion

to measure its concentration when the waveguide was brought in contact with a solution of CaCl_2 . The absorption characteristics of the doped PVC waveguide in the range from 500 to 700 nm are changed by the presence of Ca^{2+} , causing a reduction in absorption and a shift of the peak to shorter wavelength. These changes can be calibrated with respect to calcium concentration. The use of different dopants can be expected to make the waveguide sensitive to the ions of other elements.

20.2 Opto-Electronic Integrated Circuits

In the optical integrated circuits described in Section 20.1 all of the key elements were optical devices. However, there is another class of optical integrated circuits in which many of the devices are purely electronic and the signal is carried in parts of the circuit by electrical voltage or current waves rather than by an optical beam. Such circuits have come to be called opto-electronic integrated circuits (OEIC's). They are usually fabricated on substrates of semi-insulating GaAs or InP because both electronic and optical devices can be monolithically integrated on these materials. However, a silicon substrate also can be used when a hybrid approach is used, with light emitters being made in a III-V material.

20.2.1 An OEIC Transmitter

An OEIC four-channel optical transmitter is shown in Fig. 20.10 [33]. This circuit was fabricated on a semi-insulating GaAs substrate by molecular beam epitaxy. It features an array of four stripe-geometry, single-quantum-well GaAlAs lasers with microcleaved facets. Each laser is accompanied by a photodiode to monitor output power and a driver circuit containing three field effect transistors (FET's). The lasers have a relatively low threshold current of 15–29 mA and a differential quantum efficiency of 50–60%. The emission wavelength is 834 nm. The FET's are Schottky-barrier-gate devices. By monolithically integrating electronic and optical devices in an OEIC such as this, one can reduce parasitic capacitance and inductance to a minimum, thereby increasing the achievable maximum frequency of operation. In the circuit shown in Fig. 20.10 the transmitter is capable of operating at a data rate of 1.5 Gb/s. Matsueda and Nakamura [34] have produced a single-channel OEIC transmitter featuring a four FET driver circuit, a monitor photodiode and a GaAlAs laser diode, all monolithically integrated on a GaAs semiinsulating substrate. That OEIC is capable of operating at a data rate of 2 Gb/s. Other OEIC transmitters are described in a review paper by Matsueda [35].

A four-channel laser transmitter OEIC for operation at $1.55 \mu\text{m}$ wavelength has been produced by Woolnough et al. [36]. It incorporates on an InP substrate, four single-mode, ridge-waveguide, InGaAsP, DFB lasers operating in the 1545 to 1560 nm window with 0.8 nm separation between adjacent channels. The drive

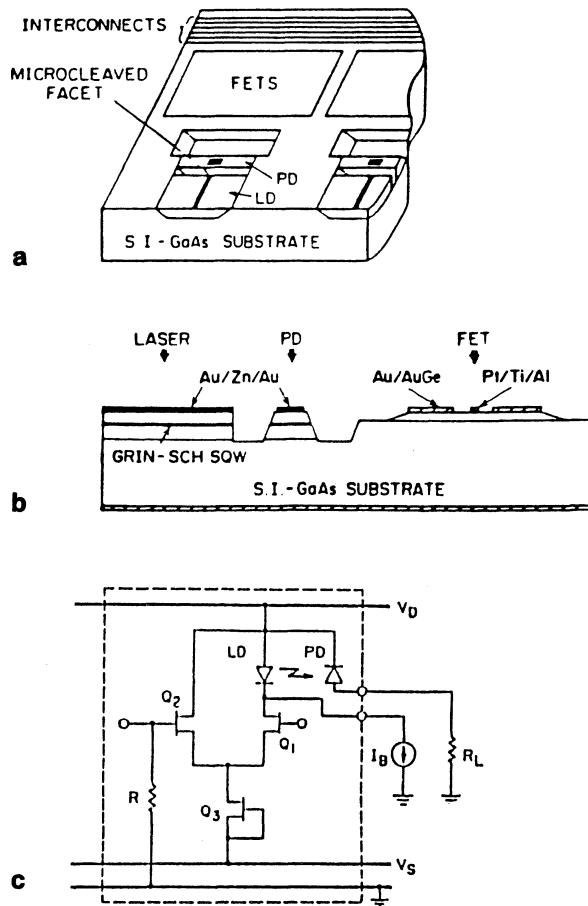


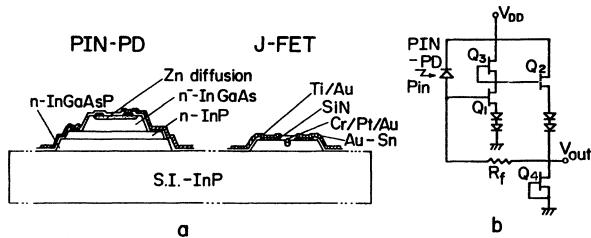
Fig. 20.10 a,b,c Schematic of four-channel OEIC transmitter **a** Over-all view showing the layout, **b** Cross-section along the laser cavity showing the device structure, **c** Circuit diagram of OEIC transmitter. The *dashed line* shows the part of monolithic integration for a single channel

circuits consist of lattice-matched, diffused-junction InGaAs-channel JFETs. Laser modulation at 155 Mbit/s has been achieved.

20.2.2 An OEIC Receiver

The detection and amplification functions of a receiver can also be implemented in OEIC form. Figure 20.11 shows a typical OEIC receiver [37]. The circuit features a p-i-n photodiode detector and Schottky-barrier gate FET amplifier monolithically integrated on a semi-insulating InP substrate. It is designed to operate with a 5 volt

Fig. 20.11 a,b OEIC photoreceiver a. Cross section of circuit chip; b. Circuit diagram [37]



power supply, which simplifies its interconnection with standard 5 volt logic IC's. The circuit has a 3 dB bandwidth of 240 MHz, with a 965Ω transimpedance.

The increased speed of OEIC's as compared to that of discrete element circuits makes them attractive for use in lightwave communication systems, signal processing and sensing applications. Data rates in excess of 40 Gb/s have been achieved in an InP-based OEIC receiver that incorporated a waveguide-integrated photodiode and a distributed amplifier that contained 4 high-electron-mobility transistors (HEMTs) [38]. Optical transceivers that combine both the transmitter and receiver functions also can be made in OEIC form [39–41].

Overall data rates of 100 Gb/s have been achieved in OEIC transceivers. For example, Kish et al. [41] have produced a monolithically integrated chip in InP that features 10 separate transceivers, each operating at a 10 Gb/s data rate. Their outputs are dense-wavelength-division-multiplexed (DWDM) onto a single optical fiber.

20.2.3 An OEIC Phased-Array Antenna Driver

Phased-array antennas have been used for many years in microwave applications in which a scanning microwave beam is required, but a moving antenna structure is impractical [42]. For example, the antenna for a RADAR transmitter in a supersonic aircraft can not be effectively scanned mechanically at a rapid enough rate. In this case, an electronically scanned phased-array antenna is used, consisting of a relatively large number of emitting antenna elements spaced many wavelengths apart, usually along the wings of the aircraft. If the relative phases of the waves transmitted by the various antenna elements are properly adjusted, a scanning microwave beam can be produced. To maintain phase coherence a frequency reference is provided by a stable master oscillator. This reference signal must be properly phase shifted and conveyed to each of the transmitting elements. If metallic microwave waveguide or coaxial cable is used to transport this signal, a great deal of undesirable weight is added to the aircraft. Obviously, this weight can be eliminated if the phase control signals are converted to optical signals and are distributed to the microwave emitters via optical fibers [43].

An OEIC which generates such phase control signals is shown in Fig. 20.12 [44]. The phase control signals are generated in a GaAs monolithic microwave integrated circuit (MMIC) with an integrated laser and driver circuit. These signals are then

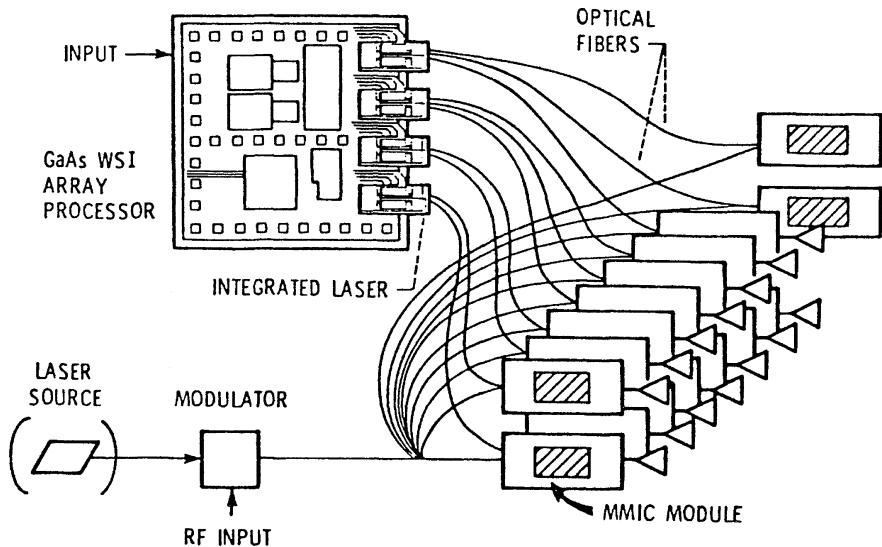


Fig. 20.12 OEIC phased array antenna driver [44]

carried by optical fibers to the microwave generating MMIC modules which are the antenna elements. In the system shown in Fig. 20.12, which was designed for a data transmission application rather than RADAR, an information signal produced by modulating a laser was also carried to all of the antenna elements by optical fiber links.

20.3 Devices and Systems for Telecommunications

While optical integrated circuits are mostly emerging from the laboratory to become commercial products, optical fibers have already found widespread use in telecommunication systems. Almost from its beginning, the use of fiber waveguides in telecommunication systems has been an international phenomenon. Fiber links have been used in many countries to transmit audio and video signals, as well as digital data. There are presently so many examples of fiber telecommunication systems that have either been already implemented or are in the final planning stage that it is impractical to mention all of them in this chapter. Consequently, only a number of representative systems will be considered.

20.3.1 Trends in Optical Telecommunications

Over the past twenty-five years there has been remarkable improvement in the capabilities of optical fiber communications systems. The systems implemented before

Table 20.3 Commercial optical fiber communications links (pre-1980)

Company	Location	Length	Performance data
AT&T	Atlanta	10.9 km	44.7 Mbit/s 144 fibers, 6.2 dB/km
GTE	Long Beach	9 km	1.5 MHz 6 fibers, 6.2 dB/km
ITT-STL	Harlow	9 km	140 Mbit/s 4 fibers, 5 dB/km
Teleprompter	So. Cal.	240 m	CATV trunk 10 K subscribers
Rediffusion Ltd.	Hastings	1.4 km	CATV trunk 34 K subscribers
AT&T	Chicago	2.5 km	44.7 Mbit/s 12 fibers 8.5 dB/km
British Telcom	Brownhills-Walsall Croydon-Vauxhall London-Vauxhall	Total length 28 km	8 Mbit/s
GEC	London Subway	7 km	8 Mbit/s
Philips	Eindhoven to Helmond	14 km	140 Mbit/s 12 fibers, graded index, 1920 telephone channels/fiber
Siemens	Frankfurt/Main Oberursel	15.4 km	34 Mbit/s
Thomson-CSF	Paris	7 km	34 Mbit/s, 50 fibers, graded index 30 000 telephone channels
Martin Marietta Data Systems	Orlando, Fla.	9.2 km	45 Mbit/s
Israeli Post Office (Fibronics Fibers)	Tel-Aviv	2.7 km	8 fibers, 148 MHz, 6 dB/km loss at 0.82 μm

approximately 1980, listed in Table 20.3, generally had a length on the order of 10 km, operated at data rates less than 150 Mb/s and carried at most 30 000 telephone channels. By contrast, the systems which became operational after about 1980, shown in Table 20.4, had lengths of thousands of kilometers, data rates ranging from hundred of megabits per second to gigabits per second and carried as many as 40 000 telephone channels. Also the spacing of repeaters increased from a few kilometers in early systems to over 100 km.

After 1990, as shown in Table 20.5, improvements in technology have continued, expanding system capacity to up to 480 000 telephone channels through the use of dispersion shifted fiber and dense wavelength division multiplexing (DWDM). The use of erbium doped fiber amplifiers (EDFAs) has permitted repeater spacings of as much as 400 km and overall system lengths on the order of 10 000 km. These improvements in the performance of telecom/datacom systems over the past 25 years have come about because of a number of key technological breakthroughs.

Table 20.4 Commercial optical fiber communications links (post-1980)

Company	Location	Length	Performance data
Pacific Telephone Co.	Sacramento, San Jose Stockton, Oakland San Francisco	Total length 270 km	
AT&T	Cambridge, MA Richmond, VA	1250 km	wavelength-division multiplexed 270 Mbit/s
AT&T	California Coast	830 km	wavelength-division multiplexed 270 Mbit/s
AT&T	Atlanta (1982)	65 km	432 Mbit/s, 6048 channels WDM (1335 & 1275 nm) single mode fiber, no repeater
AT&T	Atlanta (1984)	74 mi	420 Mbit/s, 6000 channels single mode laser, 1500 nm single mode fiber, no repeater
AT&T	5 intercity links Phila.-Pittsburgh Pittsburgh-Cleveland Dallas-Houston San Antonio-Segua Atlanta-Charlotte		432 Mbit/s, 6048 channels single mode laser, 1300 nm two single mode fibers (no WDM) repeater spacing 46 ml
NTT (F400-M) commercial	Asahikawa to Kagoshima (main trunk)	4000 km	1.3 μm, single mode 40 Km repeater spacing 445 Mbit/s, Ge APD completion 1985
Field trial			1.55 μm, single mode InGaAs APD 80-120 km repeater spacing 5760 telephone channels
NTT (F1.6G) Field Trial		4000 km	1.6 Gbit/s DFB laser 1.3 μm-40 km re. sp. 1.55 μm-80-120 km re. sp. 23 040 telephone channels
NTT (FS-400 M)		1000 km (2 sections)	Submarine cable
Telecom Australia	Perth to Adelaide	2800 km	565 Mbit/s single mode 1.3 and 1.55 μm completion 1989
International Consortium- (TAT-8)	New Jersey to Widemouth, UK Penmarch, France	6687 km	Submarine cable 40 000 telephone channels 55 km repeater spacing single mode, 1.3 μm 274 Mbit/s completion 1988
International	North America		submarine cable

Table 20.5 Commercial optical fiber communications links (post-1990)

Company	Location	Length	Performance data
International Consortium of 25 Organizations (TAT-9)	North America to Europe & U.K.		80 000 telephone channels submarine cable 565 Mbit/s completion 1991
Consortium of 4 Organizations (Taino-Carib)	Puerto Rico St. Thomas, U.S. V.I. Tortola, Brit. V.I.	175 km	225 000 telephone channels submarine cable WDM, 565 Mbit/s no repeater completion 1993
MCI	MCI backbone network	total 58 000 route kilometers	Up to 400 000 telephone channels wavelength division multiplexing dispersion shifted fiber 400 km repeater spacing (3 EDFA's in link) completion 1996
International Consortium of 30 Organizations (Columbus III)	Southern Europe to United States	10 000 km	2 fiber pairs 8 wavelengths/pair 2.5 Gbit/s aggregate of 40 Gb/s 480 000 telephone channels completion 1999
International Consortium of 50 Organizations (TAT-14)	United States to the United Kingdom, France, The Netherlands, Germany, and Denmark.	15 000 km	DWDM with 16 wavelengths, 640 Gb/s, 2 service fiber pairs, completion 2001
Joint venture between Cable & Wireless and Alcatel (Apollo)	Northern leg: Bude (UK) to Shirley, Long Island. Southern leg: Lannion (France) to Manasquan, New Jersey.	(approx. 6 000 km) (approx. 6 315 km)	DWDM 4 fiber pairs, 3.2 Tb/s each leg, completion 2003

Perhaps the most significant breakthrough was the development of efficient GaInAsP lasers and photodiodes operating at either 1.3- μm or 1.55- μm wavelength to take advantage of low loss, minimal dispersion optical fibers which were also developed for those wavelengths. As mentioned in Chapter 1, losses of less than 0.4 dB/km at 1.3- μm and 0.2 dB/km at 1.55- μm can be obtained. Minimal bulk

dispersion can also be obtained either at $1.3\text{-}\mu\text{m}$ wavelength, or at $1.55\text{-}\mu\text{m}$ in “dispersion shifted” fibers.

The serious problem of modal dispersion which limited the distance-bandwidth product of early multimode systems has been completely eliminated in modern single-mode systems. The transition from multimode to single-mode systems required the development of new single-mode fibers, couplers and lasers, and did not happen overnight. There was a time, not very long ago, when many people felt the problems of single-mode coupling, with its submicrometer tolerances, could not be overcome. Fortunately, they were wrong. Single-mode couplers which can be installed by technicians and repairmen “in the field” are now commercially available from many different suppliers. For example, the AT&T Single-Mode ST^R connector, shown in Fig. 20.13, is a low-loss, low-reflection connector suitable for high-speed, wideband digital transmission [45]. The insertion loss of this connector average 0.34 dB with a standard deviation of 0.28. Reflection averages -42.8 dB with a worst case of -34.1 dB. Practical single-mode optical fiber cables and lasers are also commercially available for use in telecommunications systems [46].

In optical fiber communication systems operating at data rates greater than 1 Gb/s, generally DFB lasers have come to be used. The very narrow emission linewidth of these lasers ($< 1 \text{ \AA}$) results in the least bulk dispersion effect and provides distance-bandwidth products in excess of 2000 GHz km when the best fiber is also used. For example, $1.55 \mu\text{m}$ wavelength transmission has been reported in laboratory tests at 20 Gb/s over 109 km [47]. In general when high-power laser diodes are modulated at high data rates, frequency shift or “chirping” due to index change from carrier injection or temperature variation can be a problem. However, the inherent frequency selectivity of the DFB laser tends to limit chirping. Specially designed

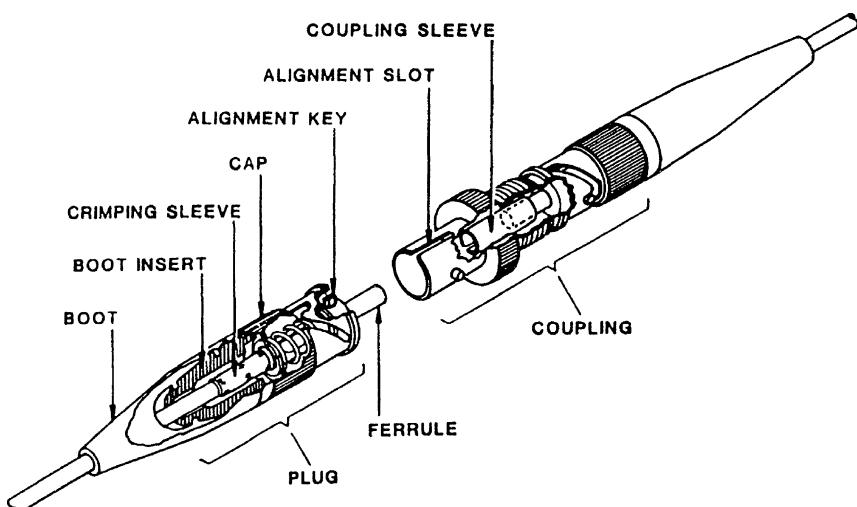


Fig. 20.13 AT&T SM ST connector [45]

DFB lasers have exhibited 3 dB modulation bandwidth in excess of 13 GHz, with no-significant chirping being reported [48].

The narrow linewidth and frequency stability of the DFB laser make it particularly useful in wavelength-division-multiplexed systems. A large number of information signal channels can be transmitted by using an array of DFB lasers, each operating at a slightly different wavelength. Typical wavelength spacing between channels is ten or twenty Angstroms; thus, many channels can be fit into the low-loss, low-dispersion passband of the optical fiber. Wavelength-division-multiplexing has been used in a number of the systems listed in Table 20.5 to achieve increased transmission capacity.

In recent years the combination of advanced DFB lasers, very-low-loss fibers, GHz-data-rate modulators and EDFA's has led to DWDM systems in which overall terabit/s data rates are transmitted over hundreds of km [49–51]. For a review of the development of optical communications systems, see e.g. Kogelnik [52].

20.3.2 New Devices for Telecommunications

In addition to the major breakthroughs described in the previous section, there are also a number of new devices and techniques that are having an effect on telecommunications and offer great promise for the future. Certainly the quantum well devices described in Chapter 18, have improved the performance of telecommunication systems. Also, optical amplifiers such as those described in Chapter 13 are finding implementation in newly installed systems. Such amplifiers can increase system bandwidth by eliminating the necessity of converting the optical signal to an electrical one for amplification in repeaters. OEIC receivers, transmitters and transceivers such as those described in Section 20.2 have increased data rates per channel.

An array waveguide (AWG), such as the device that is discussed in Section 5.4.1, can be used to make a optical add drop multiplexer (OADM). The OADM functions to couple channels containing signals either onto or off of a dense wavelength division multiplexed (DWDM) optical fiber transmission line at desired points along its length. The basic operation of an OADM is described in Section 19.4. These devices are key elements of modern fiber-optic telecom/datacom systems. The operation of an OADM that uses a single AWG to simultaneously perform both the multiplexing and demultiplexing functions has been analyzed by Gemelos et al. [53]. For a device of this type with N channels, consisting of m add/drop channels and n pass channels, the performance of the OADM degrades as either N or m is increased because of crosstalk in the AWG. Thus the number of channels in the network is limited. However, when a suitable optical filter is added at the output to remove crosstalk, the digital signal-to-noise ratio can be made to be independent of N and m .

Improvements in optical fibers have also contributed to the better performance of optical communications systems. OH-depleted single-mode fibers have broadened the range of usable wavelengths in the third window so that it now extends from

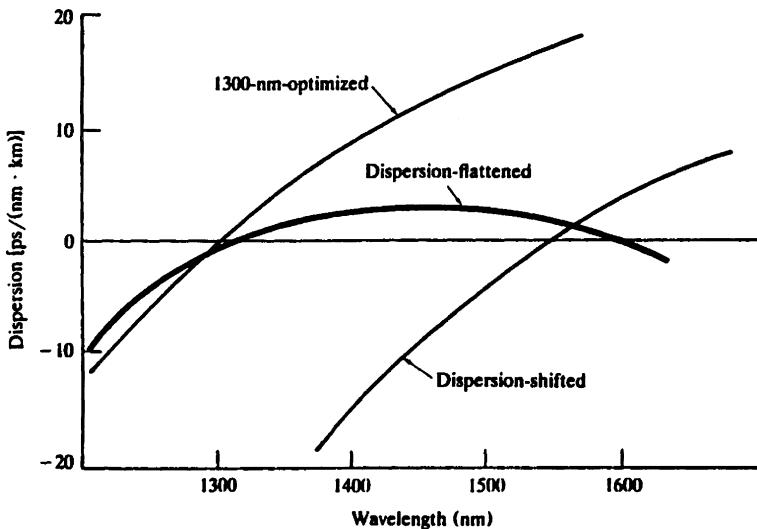


Fig. 20.14 Dispersion in dispersion-shifted and dispersion-flattened fibers

about 1450 to 1600 nm (Chapter 13 [4]). This increases the number of channels possible in DWDM systems. Dispersion shifted single-mode fibers permit the effect of dispersion to be minimized at 1.55 μm wavelength, and dispersion-compensated (flattened) single-mode fibers allow the same fiber to be used with minimum dispersion at either 1.3 or 1.55 μm , as shown in Fig. 20.14.

These improved fibers, coupled with EDFAs and high-data-rate laser/modulators, have led to systems with data-rate distance products on the order of millions of Gb/s \cdot km [49–51, 54].

One new device for optical communications that has recently emerged from the laboratory to become a commercial product is the VCSEL array. Because they emit vertically through the surface with a narrow emission angle rather than through an endface, and because they can be shaped to have a circular cross-section, VCSELs are ideally suited for efficiently coupling to optical fibers. If they are integrated into an array, perhaps including electronic driving circuitry to form an OEIC, they offer a very convenient source for multifiber systems. An example of such a source is the two-dimensional VCSEL array for high-speed parallel links reported by Kosaka [55]. The details of his array are shown in Fig. 20.15.

All epitaxial layers are grown on an n-GaAs substrate by molecular beam epitaxy (MBE). Three 0.01 μm thick InGaAs strained multiple-quantum-well (MQW) layers, serving as active layers, are placed at the peak of the electric-field standing wave inside one-wave AlGaAs cladding layers. The composition ratio of each InGaAs layer is different to obtain different bandgap wavelengths of 990, 980, and 970 nm which enables a broad-gain bandwidth. The cladding layers are sandwiched between two distributed Bragg reflectors (DBR's) consisting of GaAs-AlAs quarter-wave multiple stacks with linearly graded superlattice transition layers. The VCSELs had

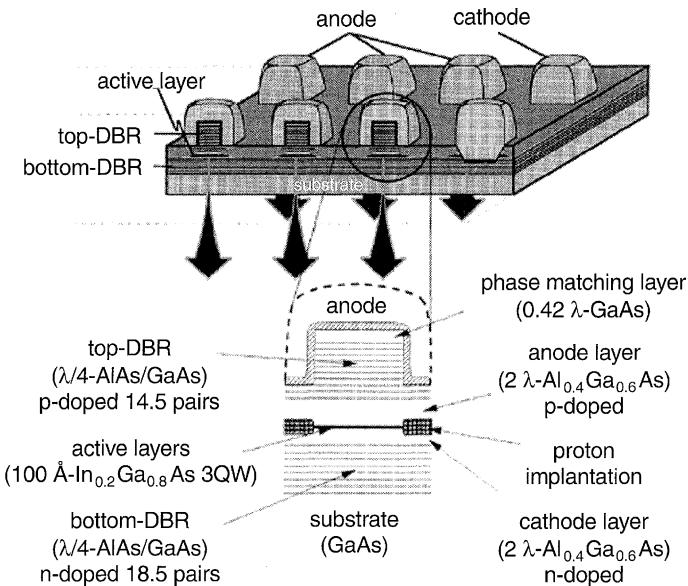


Fig. 20.15 VCSEL Array [55] © 1999 IEEE

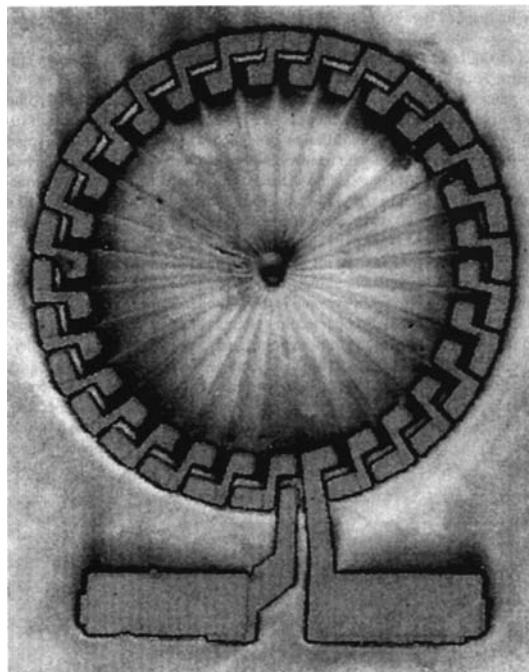


Fig. 20.16 Photomicrograph of 30 sector photogenerator (approx. 1 mm diameter)

a threshold current of 1.5 mA and a quantum efficiency of 0.29 W/A. When the 8×8 array was mounted in an individually addressable, 64 channel, strip-line-fed microwave package, it exhibited a 3 dB modulation bandwidth of 7 GHz.

Another device that is just emerging from the laboratory, but may have an important impact on future telecom systems is the photogenerator made by Dentai et al. [56]. In metal-wire electronic telecom systems electrical power can be transmitted in the wires along with the information signals, so that switches, controllers and electronic circuits can operate without dependence on the electrical power grid. The lack of this capability has for some time seemed to be an inherent problem of optical fiber systems. This new photogenerator solves that problem by generating power at voltages in excess of 10 V using only 1480–1650 nm light that can be transmitted through the fibers of an optical communication system. The optical power is converted to electrical power by InGaAs photodiodes grown on an InP substrate. Each diode generates less than 0.5 V, but they can be connected in series to generate higher voltages. Figure 20.16 shows a 30 sector photogenerator that produces 10 V.

When illuminated with $500 \mu\text{W}$ of 1554 nm light, this photogenerator produced 10.3 V. It had a short-circuit responsivity of 0.025 A/W . This is sufficient to power a variety of lightwave circuits and MEMs.

Problems

- 20.1 What are the key technological developments over the past few years that have led to telecom/datacom systems with overall terabit per second data rates?
- 20.2 What is the purpose of an OADM?
- 20.3 What is dispersion shifted fiber?
- 20.4 List three different types of sensors that can be made in integrated optic form.
- 20.5 During what decade were the first optical fiber communications links established?
- 20.6 What type of laser is used in a Blu-ray Disc drive?
- 20.7 What is dense wavelength division multiplexing (DWDM)?
- 20.8 What are the advantages of optical sensors as compared to electrical (electronic) sensors?

References

1. M.C. Hamilton, D.A. Wille, W.J. Micelli: Opt. Eng. **16**, 475 (1977)
2. D. Mergerian, E.C. Malarkey: Microwave J. **23**, 37 (May 1980)
3. D. Mergerian, E.C. Malarkey, R.P. Pautienus, J.C. Bradley, M. Mill, C.W. Baugh, A.L. Kellner, M. Mentzer: SPIE Proc **321**, 149 (1982)
4. A.F. Milton: Charge transfer devices for infrared imaging, in *Optical and Infrared Detectors*, R.J. Keyes 2nd edn., Topics Appl. Phys., Vol. 19 (Springer, Berlin, Heidelberg 1980)
5. D.F. Barbe (ed.): *Charge-Coupled Devices*. Topics Appl. Phys., Vol. 38 (Springer, Berlin, Heidelberg 1980)

6. M.E. Pedinoff, T.R. Ranganath, T.R. Joseph, J.Y. Lee: NASA Conf. on Optical Information Processing for Aerospace Applications, Houston, TX (1981) Proc. p. 173
7. B. Chen, T.R. Joseph, J.Y. Lee: IGWO'80, Incline Village, NV Paper ME-3
8. K. Aiki, M. Nakamura, J. Umeda: IEEE J. **QE-13**, 220 (1977)
9. K. Aiki, M. Nakamura, J. Umeda: IEEE J. **QE-13**, 597 (1977)
10. C. Zah, R. Bhat, B. Pathak, L. Curtis, F. Favire, P. Lin, C. Caneny, A. Gozdz, W. Lin, N. Andreadakis, D. Mahoney, M. Koza, W. Young, T. Lee: IEEE 5th Int'l Conf. on InP and Related Materials, Paris (1993) Proc. p. 77
11. K. Sato, S. Sakine, Y. Kondo, M. Yamamoto: IEEE J. **QE-29**, 1805 (1993)
12. S.L. Lee, I.F. Jang, C.Y. Wang, C.T. Pien, T.T. Shih: Monolithically integrated multiwavelength sampled grating DBR lasers for dense WDM applications. IEEE J. Sel. Top. Quant. Electron. **6**, 197 (2000)
13. H.F. Taylor: IEEE J. **15**, 210 (1979)
14. H.F. Taylor, M.J. Taylor, P.W. Bauer: Appl. Phys. Lett. **32**, 559 (1978)
15. S. Yamada, M. Minakota, J. Noda: Appl. Phys. Lett. **39**, 124 (1981)
16. M.J. Demler: *High-Speed Analog-to-Digital Conversion* (Academic Press, San Diego, CA 1991)
17. W-R. Yang, T. Renkoski, W. Nunnally: Overview of an all-optical analog-to-digital converter: focusing of deflected terahertz optical pulse and propagation characteristics, Proc. SPIE **5814**, 62 (2005)
18. J.U. Kang, R.D. Esman: Demonstration of time interwoven photonic four-channel WDM sampler for hybrid analogue-digital converter. Electron. Lett. **35**, 60 (1999)
19. M. Currie, T.R. Clark, P.J. Matthews: Photonic analog-to-digital conversion by distributed phase modulation. IEEE Photonics Tech. Lett. **12**, 1689 (2000)
20. P. Rabiei, A.F.J. Levi: Analysis of hybrid optoelectronic WDM ADC. IEEE J. Lightwave Tech. **18**, 1264 (2000)
21. J.C. Twichell, R. Helkey: Phase-encoded optical sampling for analog-to-digital converters. IEEE Photonics Tech. Lett. **12**, 1237 (2000)
22. H. Toda, M. Haruna, N. Nishihara: IEEE J. **LT-5**, 901 (1987)
23. H. Toda, K. Kasazum, M. Haruna, N. Nishihara: IEEE J. LT-7, 364 (1989)
24. S. Ura, T. Suhara, H. Nishihara: IEEE J. **LT-4**, 913 (1986)
25. S. Ura, M. Shimohara, T. Suhara, N. Nishihara: IEEE Photon. Tech. Lett. **6**, 239 (1994)
26. J.A. Hudgings, S.F. Lim, G.S. Li, Wupen Yuen, K.Y. Lau, C.J. Chang-Hasnain: Compact, intergrated optical disk readout head using a novel bistable vertical-cavity surface-emitting laser. IEEE Photonics Tech. Lett. **11**, 245 (Feb. 1999)
27. K. Manoh, H. Yoshida, T. Kobayashi, M. Takase, K. Yamauchi, S. Fujiwara, T. Ohno, N. Nishi, M. Ozawa, M. Ikeda, T. Tojyo, T. Taniguchi: Small integrated optical head device using a blue-violet laser diode for Blu-ray Disc system, Jpn. J. Appl. Phys. **42**, 880 (2003)
28. L. Johnson, F. Leonberger, G. Pratt: Appl. Phys. Lett. **41**, 134 (1982)
29. N.A.F. Jaeger, L. Young: IEEE J. **LT-7**, 229 (1989)
30. R.F. Nabiev, C.J. Chang-Hasnain, L.E. Eng: Spectrodetector-Novel Monolithic wavelength meter and photodetector. Electron. Lett. **31**, 1373 (1995)
31. C.K. Madsen, J. Wagener, T.A. Strasser, D. Muehlner, M.A. Milbrodt, E.J. Laskowski, J. DeMarco: Planar waveguide optical spectrum analyzer using a UV enhanced grating. IEEE J. Sel. Top. Quant. Elect. **4**, 925 (1998)
32. K. Kim, H. Minamitani, K. Matsumoto, S. Kang: Sensing property of thin film optical waveguide sensor based on PVC co-polymer. Proc. SPIE **3278**, 220 (1998)
33. O. Wada, H. Nobubara, T. Sanada, M. Kuno, M. Makiuchi, T. Fujii, T. Sakurai: IEEE J. **LT-7**, 186 (1989)
34. H. Matsueda, M. Nakamura: Appl. Opt. **23**, 779 (1984)
35. H. Matsueda: IEEE J. **LT-5**, 1382 (1987)
36. P. Woolnough, P. Birdsall, P. O'Sullivan, A. Cockburn, M. Harlow: Electron. Lett. **29**, 1388 (1993)

37. T. Horimatsu, M. Saskai: IEEE J. LT-7, 1612 (1989)
38. G.G. Mekonnen, W. Schlaak, H-G. Bach, R. Steingruber, A. Seeger, Th. Enger, W. Passenger, A. Umbach, C. Schramm, G. Unterborsch, S. van Waasen: 37 GHz bandwidth InP-based photoreceiver OEIC suitable for data rates up to 50 Gb/s. IEEE Photonics Tech. Lett. **11**, 257 (1999)
39. D.A. Louderback, O. Sjolund, E.R. Hegblom, S. Nakagawa, J. Ko, L.A. Coldren: Modulation and free-space link characteristics of monolithically integrated vertical-cavity lasers and photodetectors with microlenses. IEEE J. Sel. Top. Quant. Electron. **5** 157 (1999)
40. K. Kato, Y. Thomori: PLC hybrid integration technology and its application to photonic components. IEEE J. Sel. Top. Quant. Electron. **6**, 4 (2000)
41. F.A. Kish, R. Nagarajan, C.H. Joyner, R.P. Schneider, Jr., J.S. Bostak, T. Butrie, A.G. Dentai, V.G. Dominic, P.W. Evans, M. Kato, M. Kauffman, D.J.H. Lambert, S.K. Mathis, A. Mathur, R.H. Miles, M.L. Mitchell, M.J. Missey, S. Murthy, A.C. Nilsson, F.H. Peters, S.C. Pennypacker, J.L. Pleumeekers, R.A. Salvatore, R.S.G. Reffle, D.G. Mehuy, D. Perkins, D.F. Welch: 100 Gb/s (10 x 10 Gb/s) DWDM photonic integrated circuit transmitters and receivers, Lasers and Proceedings of Conference on Electro-Optics, 2005. (CLEO). **1**, 585 (2005)
42. J.-F. Luy, P. Russer (eds.): *Silicon-Based Millimeter-Wave Sereives*, Springer Scr. Electron. Photon., Vol. 32 (Springer, Berlin, Heidelberg 1994)
43. R.G. Hunsperger, M.K. Barnoski, H.W. Yen: A system for optical injection locking and switching of microwave oscillators. US Patent No. 4,264,857 (issued April 1981)
44. R.R. Kunath, K.B. Bhasin: IEEE AP-S/VRSI Symp., Philadelphia, PA (1986)
45. G.M. Alameel, A.W. Carlisle: Fiber and Integ. Opt. **8**, 45 (1989)
46. Laser Focus World Buyers Guide '01 (PennWell Publishing, Tulsa, OK 2001)
47. D. Mathoorasing, C. Kazmiricrski, M. Blez, Y. Sorel, J. Kerdiles, M. Henry, C. The bault: Electron. Lett. **30**, 507 (1994)
48. Y. Hirayama, H. Furuyama, M. Moringa, N. Suzuki, M. Kushibe, K. Eguchi, M. Nakamura: IEEE J. QE-25, 1320 (1989)
49. H. Onaka, H. Miyata, G. Ishikawa, K. Otskula, H. Ooi, Y. Kai, S. Kinoshita, M. Seino, H. Nishimoto, T. Chikarna: 1.1 Tb/s WDM transmission over a 150 km 1.3 μ m zero-dispersion single-mode fiber. Proc. Opt. Fiber Conf., San Jose, CA, 1996
50. A.H. Gnauck, A.R. Chraplyvy, R.W. Tkach, J.L. Zyskind, J.W. Sulhoff, J. Lucero, Y. Sun, R.M. Jopson, F. Forghieri, R.M. Derosier, C. Wolf, A.R. McCormick: One terabit/s transmission experiment. Proc. Opt. Fiber Conf., San Jose, CA, 1996
51. T. Morioka, H. Takara, S. Kawanishi, O. Kamatani, K. Takiguchi, K. Kuchiyama, M. Saruwatari, H. Takahashi, M. Yamada, T. Kanamori, H. Ono: 100 Gbit/s 10 channel OTDM/WDM transmission using a single supercontinuum WDM source. Proc. Opt. Fiber Conf., San Jose, CA, 1996
52. H. Kogelnik: High-capacity optical communications: personal recollections. IEEE J. Sel. Top. Quant. Electron. **6**, 1279 (2000)
53. S.M. Gemelos, D. Wonglumsorm, L.G. Kazovsky: Impact of crosstalk in an arrayed-waveguide router on an optical add-drop multiplexer. IEEE Photonics Tech. Lett. **11**, 349 (1999)
54. K. Imai, T. Tsuritani, N. Takeda, K. Tanaka, N. Edagawa, M. Suzuki: 500 Gb/s (50 \times 10.66 Gb/s) WDM transmission over 4000 km using broad-band EDFA's and low dispersion slope fiber. IEEE Photonics Tech. Lett. **12**, 909 (2000)
55. H. Kosaka: Smart integration and packaging of 2D VCSEL's for high-speed parallel links. IEEE J. Sel. Top. Quant. Elect. **5**, 184 (1999)
56. A.G. Dentai, C.R. Giles, E. Burrows, C.A. Burrus, L. Stulz, J. Centanni, J. Hoffman, B. Moyer: A long-wavelength 10-V optical-to-electrical InGaAs photogenerator. IEEE Photonics Tech. Lett. **11**, 114 (1999)

Chapter 21

Photonic and Microwave Wireless Systems

Beginning in about 1990, a new trend became significant in the field of integrated optics; it is the merging of photonic devices and systems with those of RF (radio frequency) and microwaves. Our telecom/datacom systems mostly now have been converted to lightwave systems, employing optical fibers and integrated optic devices. These systems have vastly increased the capacity and effectiveness of worldwide communications networks, as discussed in Chapter 20. However, they leave still a missing link to the end user. We are a mobile society and it isn't always convenient to be connected to the network by a glass fiber or a metal wire. This has led to the development of "wireless" radio communications networks based on the transmission of RF or microwave signals through the air. However, it is not practical for everyone to carry a high-power radio transceiver. By combining short-distance, low-power wireless system with lightwave systems employing optical fibers and OICs for long distance transmission, we have made it possible, for example, for a person (or computer) traveling in a vehicle say in Los Angeles to communicate with one in a vehicle in major cities in Europe or Asia. The combination of fiber-optics for long-distance point-to-point communications and microwaves for short-range wireless communications provides high data rates, high levels of security and reliable mobile communications.

21.1 Merging of Photonics and Microwave Technology

In most major metropolitan areas, and in many smaller towns as well as, a communication system has evolved in which wireless cell phones complete the link to the subscriber by means of RF or microwave transmission of signals through the air to repeaters that couple the signals into the land-line telephone system. Generally waves with frequencies less than 1 GHz are considered to be RF, while those higher are called microwaves. Microwave frequencies above 30 GHz correspond to millimeter waves. Currently cell phones operate in either the RF range at about 900 MHz or in the microwave range at about 3 GHz. However, the research trend is clearly toward mm waves because shorter wavelengths permit smaller antennas and more compact phones. The cells are usually only a few miles square, and thus low

power transmitters of just a few mW can be used in the portable transceiver. The signal from a moving cell phone is handed off from one cell to another as it moves across the border between them. The phones automatically switch frequency if necessary to find an open channel in the new cell. The cell phone repeaters must be connected to the land-line system by means of either metallic cables or optical fibers. Very often fibers are used because of the advantages discussed in Chapter 1. Ultimately, one can envision a system in which lightwave signals are carried by fibers all of the way to miniature mm-wave repeater transceivers with built-in antennas located on tall buildings and towers. In a mm-wave system the size of the portable transceiver can also be reduced from today's shirt-pocket size, perhaps to wrist-watch size. A system of this type was proposed as early as 1992 [1], and perhaps earlier. Techniques for modulating semiconductor diode lasers at mm-wave frequencies have been developed, as discussed in Chapter 16, and mm-wave photodiodes are also known to be available [2]. Systems using devices of this type to transmit lightwave signals that are modulated at mm-wave frequencies are reviewed in Section 21.2.3. To achieve the very small dimensions desired in a mm-wave communication system the technologies of monolithic microwave integrated circuits (MMICs) and optical integrated circuits (OICs) must be combined to create opto-microwave monolithic integrated circuits (OMMICs). These are discussed in Section 21.2.2.

The merging of microwave and photonic technologies also can be observed in wireless local area networks (WLANs) in homes, businesses and wide-area Internet access systems. In many cases fibers and lightwaves are used to carry information and control signals to the microwave transceivers as described previously. In other applications the role of photonics is perhaps more subtle, involving free-space "optical wireless", the use of solar cells to power equipment or the use of photonic devices in the fabrication and testing of components of the wireless system.

There are many opportunities for the use of integrated optic techniques in the generation, transmission, detection and processing of microwave signals. The RF spectrum analyzer described in Section 20.1 and the phased array antenna driver (Section 20.2) are two examples of such applications. There are a number of reasons why it is often advantageous to use optical signals to control microwave devices. First is the isolation of the optical signal from undesired RF coupling encountered in the use of an electronic signal to control a microwave device or signal. This eliminates the problem of the control signal directly feeding through, or coupling, in an unwanted fashion to other channels. Optical control will significantly reduce the weight of the control signal transmission lines by replacing metallic waveguides with optical fibers. Also, optical control generally provides faster response because the optical signals are not slowed down by a parasitic capacitance and the electronic signal delays of coax or metallic waveguide. Some examples of optical signals being used to control microwave device include RF switching [3], phase shifting [4], optical injection locking of microwave oscillators [5] and optically pumped microwave mixer diodes [6]. A detailed discussion of these techniques is beyond the scope of this text but they have been reviewed elsewhere [7, 8].

In addition to being used to control microwave devices, integrated optic techniques can also be used to analyze microwave signals. The IO RF spectrum analyzer is an example of this being done in the frequency domain. An IO time domain RF

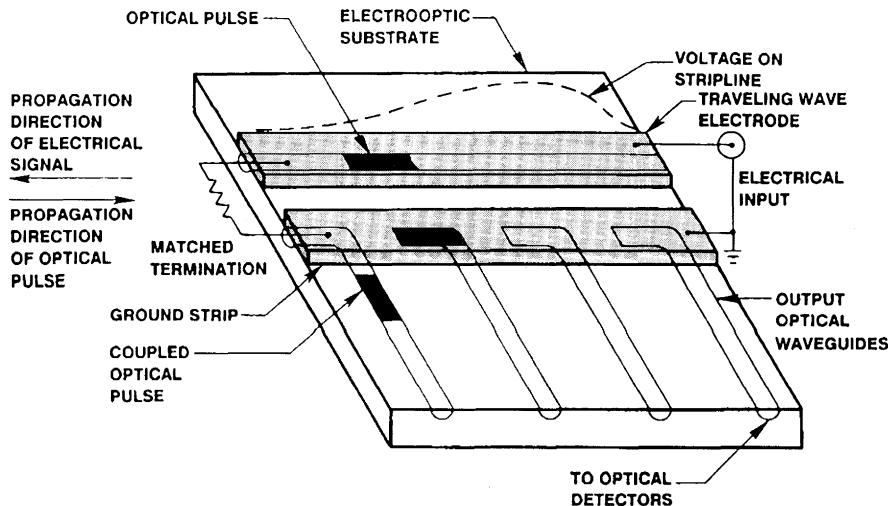


Fig. 21.1 IO high frequency waveform sampler [9]

voltage wave sampler has been produced by Ridgway and Davis [9], as shown in Fig. 21.1. This device consists of a series of electro-optic dual-channel directional couplers spaced along the length of an overlaying coplanar RF stripline which forms traveling wave electrodes for the couplers. In Operation, an optical pulse is propagated along the main waveguide in the direction opposite to the propagation of the RF voltage wave. Depending on the magnitude of the RF voltage wave at a given coupler when the optical pulse passes, a certain fraction of the light is coupled into the auxiliary waveguide of that coupler.

Thus the RF voltage waveform is sampled, and its magnitude and shape can be determined. Since the velocity of the optical pulse is much greater than that of the voltage wave, good time-domain resolution can be obtained. An optimum temporal resolution of a few picoseconds at a rate in excess of 4×10^{10} samples per second has been estimated for the OIC of Fig. 21.1, which was fabricated on a LiNbO_3 substrate with Ti diffused waveguides. The light source used was a GaAlAs laser emitting at a wavelength of $0.84 \mu\text{m}$.

While there is a fairly broad range of opto-microwave applications the one with the most widespread use has been the transmission of signals modulated at microwave frequencies over high-speed fiber optic links. The merging of photonics with RF and microwave technologies in communications has reached the point at which joint conferences can be held [10].

21.2 Fiber-Optic Transmission of RF and Microwave Signals

The transmission of RF or microwave signals on a modulated beam of light over an optical fiber link is a desirable alternative to conventional transmission over a metallic waveguide or cable. The advantages of fiber transmission are less weight, improved

security and reduced transmission loss. It is an interesting fact that reduced loss in the fiber, as compared to that in the cable or metal waveguide, more than compensates for the coupling and conversion (electrical/optical) loss. Thus in a properly designed system there is usually an insertion gain rather than an insertion loss when a cable or waveguide is replaced with an optical fiber transmission system [11].

21.2.1 Basic Principles

Transmission of microwave signals over a fiber is a relatively straightforward process, as illustrated in the block diagrams of Fig. 21.2.

The RF or microwave signals are used to intensity modulate a semiconductor laser diode. This can be done either by directly varying the diode current in proportion to the modulating signal as in Fig. 21.2(b), or by using an external modulator as in (c). In either case, suitable optical couplers and microwave impedance matching circuits must be provided to minimize reflections and coupling loss. As discussed in Chapter 16, microwave signals are usually introduced through striplines and the external modulator (if used) should have traveling-wave electrodes. Both electroabsorption and Mach Zehnder modulators are effective in this application. The optical receiver usually consists of a p-i-n photodiode with a microwave transistor amplifier. Fiberoptic systems have been used to transmit microwave signals containing voice, data and digital video at frequencies in the mm-wave range [12]. Digital data has been transmitted at frequencies as high as 60 GHz [13, 14]. Details of these systems are given in Section 21.2.3.

A relatively thorough tutorial on the design of fiberoptic microwave links has been published by Olson [15]. The two most important characteristics of a microwave-on-fiber transmission link are its insertion gain (or loss) G_{link} and its 3 dB bandwidth $f_{3\text{dB}}$. The equations necessary to calculate these two parameters, excerpted from the tutorial, are given below. The terms that make up G_{link} are the transmitter efficiency, fiber loss, receiver efficiency and the ratio of the output to input impedances. In terms of a power gain, the link gain can be written as

$$G_{\text{link}} = \left(\frac{I_{\text{out}}}{I_{\text{in}}} \right)^2 \frac{R_{\text{out}}}{R_{\text{in}}}, \quad (21.1)$$

where I_{out} and I_{in} are the output and input currents, R_{out} is the load resistance at the receiver output and R_{in} is the input resistance of the laser transmitter. The current ratio can be expanded as

$$\frac{I_{\text{out}}}{I_{\text{in}}} = \frac{\eta_{\text{txrf}} \cdot \eta_{\text{rxrf}}}{L_{\text{opt}}}, \quad (21.2)$$

where η_{txrf} = the overall efficiency of the total transmitter in mW/mA, and η_{rxrf} = the overall efficiency of the receiver in converting optical power modulation into output current in mA/mW. L_{opt} = the optical loss of the fiber link, given by the ratio

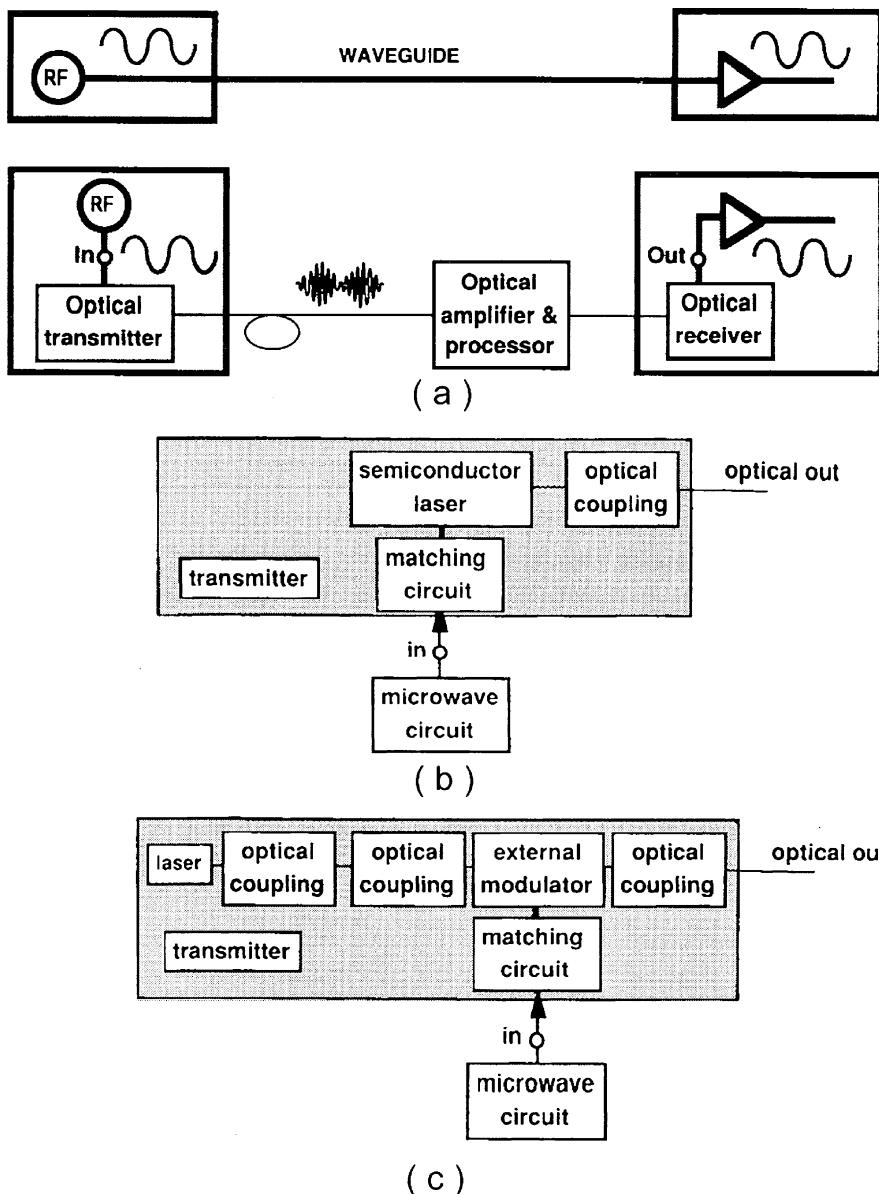


Fig. 21.2 a-c. Replacement of conventional microwave transmission system by fiberoptic link
(a) system comparison **(b)** directly modulated transmitter **(c)** external modulator Reprinted from [11] p. 280 by courtesy of Marcel Dekker, Inc.

$$L_{\text{opt}} = \frac{\text{optical power@Tx}}{\text{optical power@Rx}}. \quad (21.3)$$

Substituting (21.2) into (21.1) yields

$$G_{\text{link}} = \left(\frac{\eta_{\text{txrf}} \cdot \eta_{\text{rxrf}}}{L_{\text{opt}}} \right)^2 \frac{R_{\text{out}}}{R_{\text{in}}}. \quad (21.4)$$

In terms of dB (21.4) becomes

$$G_{\text{link}} = \underbrace{20 \log(\eta_{\text{txrf}} \cdot \eta_{\text{rxrf}})}_{\text{Electro/optic efficiency}} - \underbrace{20 \log L_{\text{opt}}}_{\text{Optical Loss}} + \underbrace{10 \log \left(\frac{R_{\text{out}}}{R_{\text{in}}} \right)}_{\text{Impedance difference}}. \quad (21.5)$$

The bandwidth of a fiberoptic microwave link is limited by the bandwidths of the transmitter and receiver and by the dispersion of the optical fiber. However it is usually the bandwidth of the transmitter that dominates. Laser transmitters typically have a relatively flat response curve up to a frequency at which the curve first peaks and then rapidly drops off. This resonance peak frequency is proportional to the square of the average optical power out of the laser P_{out} . Thus the 3 dB bandwidth of the link $f_{3\text{dB}}$ is given by

$$f_{3\text{dB}} = A(P_{\text{out}})^{1/2} \quad (2.6)$$

where A is a constant for a given diode chip

21.2.2 Device Performance

In order to produce laser transmitters and photodiode receivers that operate in the mm-wave range it is usually necessary to use either hybrid (multichip) or monolithic (single chip) integration of laser diodes with driving circuitry and photodiodes with amplifiers. The monolithic types, OMMICs, or as they are sometimes called, OEM-MICs (opto-electro-monolithic-microwave-integrated-circuits) are formed in III-V semiconductors such as GaInAsP or GaAlAs, on substrates of either GaAs or InP. These materials are very well suited to integrating microwave and optical devices on the same chip because they have the inherent properties that are desired both in OICs and MMICs. These properties are enumerated in Table 21.1

Laser/driver OMMICs are discussed in Chapter 16, so no further examples will be given in this section. An example of OMMIC receivers designed for mm-wave operation is provided by Umbach et al. [16], who have combined both p-i-n and m-s-m photodiodes with dual-gate HEMT (high-electron-mobility transistor) amplifiers. The receiver circuits were grown on Fe-doped InP (semi-insulating) substrates by MOVPE. They were designed for operation at 1.55 μm wavelength. For efficient

Table 21.1 Material properties of GaInAsP and GaAlAs desired for OICs and MMICs*Desired in OICs*

- Transparent in visible and near IR
- Direct bandgap - can make both efficient light emitters and detectors
- Energy bandgap and index of refraction can be conveniently controlled by changing composition to produce emitters, detectors and waveguides for a given wavelength

Desired in MMICs

- High electron mobility for high -gain transistors
- High scattering - limited velocity for high-frequency transistors
- Semi-insulating doping possible - can deposit microwave striplines directly on semiconductor surface.

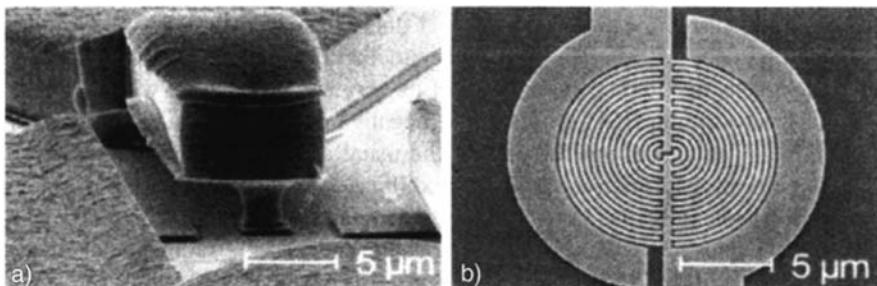


Fig. 21.3 a, b SEM photographs of photodiodes **a.** p-i-n waveguide photodiode with airbridge contact to active area **b.** top-illuminated MSM photodiode with $0.2 \mu\text{m}$ interdigitated electrodes [16] © 1999 IEEE

response to mm-wave modulation frequencies the photodiodes were required to have μm -sized dimensions, as shown in Fig. 21.3

Photodiodes of both types have been observed to have bandwidths of up to 70 GHz. They were integrated with different HEMT amplifier designs having transit frequencies as high as 90 GHz. Application of these OMMICs to a 40-Gb/s broadband photoreceiver for time division multiplexing systems and to 38-and 60-GHz narrow-band photoreceivers for use in mobile communication systems has been reported.

Integration of a p-i-n photodiode with a HEMT amplifier has also been reported by Baeyens et al. [2]. In that case an InGaAs p-i-n diode and AlGaAs/InGaAs/GaAs HEMTs were grown on a semi-insulating GaAs substrate. An air bridge contact was used to couple the photodiode to the HEMT gate to reduce capacitance. A 2-stage narrow-band amplifier based on $0.15 \mu\text{m}$ -gate, GaAs dual-gate HEMTs was used. The dual gate reduces effective gate resistance, which is a major factor in limiting high-frequency response. The photodiode was sensitive at 1.3–1.55 μm wavelengths. Two receivers were fabricated. The first, designed for operation at 42 GHz, had a responsivity of 7 A/W. The second receiver was operated at 58 GHz with a responsivity of 2.5 A/W.

21.2.3 System Performance

Quite a few microwave over fiber systems have already been implemented by using devices of the types that have been described. At mm-wave frequencies, external modulators are generally used instead of direct modulation to avoid the problem of “chirping” (frequency shift during the pulse due to the change in current flow). However, “external” doesn’t mean off of the laser chip. Generally the external intensity modulator is integrated with the laser. Both Mach Zehnder and electroabsorption modulators are used. At frequencies above about 50 GHz even the use of an external modulator cannot completely reduce chirping to insignificant levels. The modulator generates a chirped optical double sideband signal with the carrier. If this DSB signal is transmitted over a conventional fiber without correction for dispersion, chromatic dispersion sets a limit to maximum fiber length [17]. Various techniques have been developed to cope with this dispersion problem. Chirped fiber gratings can be used to compensate for dispersion [18]. A dual-electrode Mach-Zehnder modulator can be used to generate optical single sideband (SSB) modulation, which is much less sensitive to dispersion because of its very narrow spectral width [19]. An electroabsorption (EA) modulator with negative chirp has been used by Stohr et al. [17] to cancel out the effects of the usual chirp. Multiple quantum well (MQW) electroabsorption modulators can exhibit negative chirp when they are operated at wavelengths just above the exitonic resonance or when they are driven at large reverse bias level. Under those conditions the change in refractive index due to the quantum confined Stark effect (QCSE) is negative, and hence so is the chirp. (For a discussion of the QCSE see Chapter 18.) A negative-chirp EA modulator has been used to demonstrate the simultaneous transmission of two 156 Mb/s data channels in the 60-GHz band over a 1409 m length of standard single-mode fiber [17].

A full-duplex (simultaneous transmission and reception) radio-on-fiber system for the 60-GHz band, featuring an electroabsorption OMMIC optical transceiver, has been reported by Kuri et al. [20]. The OMMIC is a hybrid circuit fabricated on a dielectric substrate, as shown in Fig. 21.4.

The size of the transceiver module is 21 mm(*w*) × 13 mm(*d*) × 11.3 mm(*h*). The module has individual RF input and output ports with standard connectors. The microwaves are carrier to and from the EA diode chip by microstrip lines, and impedance matching circuits (IMC) are used to minimize reflections. Standard single-mode fibers are coupled to the EA chip through lenses. The total coupling loss was measured to be 6 dB. The temperature of the EA chip is maintained at a constant 25°C by a thermistor and built-in Peltier thermoelectric cooler (TEC). The EA chip has a 10-well InGaAsP-InGaAsP MQW core that is split into two diode regions, one for photodetection (reception) and one for EA modulation (transmission). The downlink (reception) channel was operated at 1530 nm and the uplink (transmission) was at 1570 nm. Simultaneous reception of a 59.6 GHz downlink and transmission of a 60.0 GHz uplink signal was achieved at a data rate of 156 Mb/s. A transceiver of this type has been used to transmit 60-GHz band signals over an 85 km long standard single mode fiber, using optical SSB filtering to overcome fiber dispersion [13].

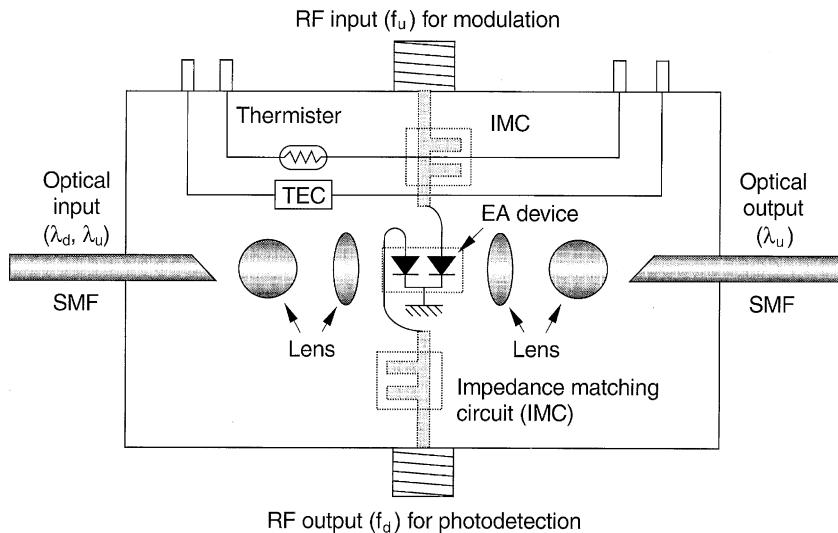


Fig. 21.4 60-GHz band EA optical transceiver, diagram of circuit [20] © 2000 IEEE

Radio-over-fiber can also be achieved by modulating a laser diode with the information signal at lower intermediate frequencies (IF) for transmission over a fiber, and then using a remote phase-locked-loop (PLL) located in a distant mm-wave radio transmitter for upconversion to mm-wave frequencies. The reference signal for the PLL is transmitted over the fiber along with the digitally modulated subcarriers that carry the signal information. A system of this type has been analyzed by Griffin et al. [21].

Transmission of RF and microwaves over a fiber on an optical beam is not limited to laboratory research and development. There are a number of companies that are presently selling radio-over-fiber links. The input to and output from these links is RF or microwaves, but the transmission is via lightwaves in a fiber. Links providing data rates of over 10 Gb/s are commercially available as off-the-shelf products.

21.3 Microwave Carrier Generation by Optical Techniques

A very interesting area of the merging of photonics and microwaves is the generation of microwaves by optical means. From the standpoint of the physicist it is a good demonstration that radio waves and lightwaves are both the same kind of electromagnetic waves, but just having different frequencies. From the standpoint of the engineer it provides a convenient way to generate required microwaves sub-carriers in frequency division multiplexed (FDM) systems. The basic concept involved in generating microwaves by using lightwaves is the well-known principle of heterodyning. When two waves of different frequency are allowed to mix together in a medium that responds to them a “beat” frequency will be generated that is equal

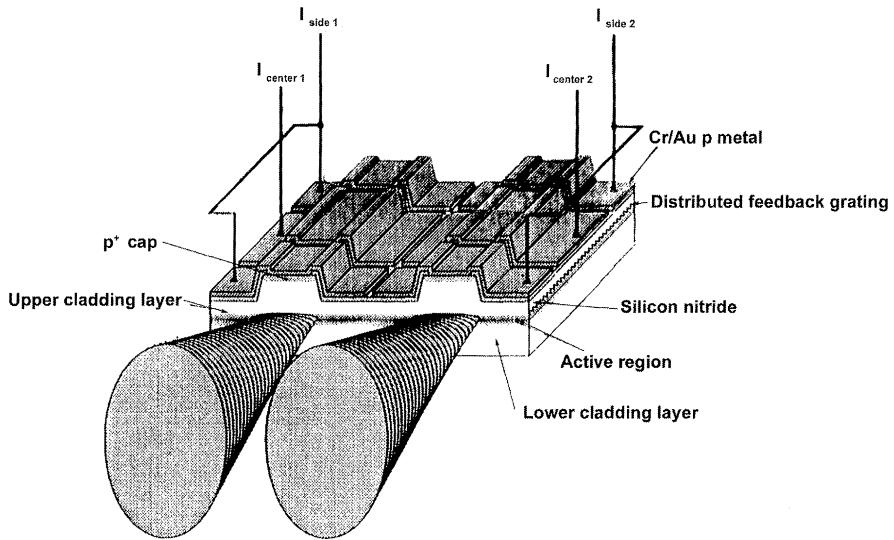


Fig. 21.5 Integrated tunable DFB lasers

to the difference between the two frequencies of the waves. This principle holds at lightwave frequencies just as well as it does at radio frequencies. Thus, for example, if the lightwaves from two tunable DFB lasers, emitting at different frequencies (wavelengths), are allowed to mix in the active region of a p-n junction photodiode, an electrical current will be generated at the beat frequency. Since the frequency of light emitted from a semiconductor laser is on the order of 10^{15} Hz, the beat frequency can easily be in the microwave range of 10^9 to 10^{11} Hz with only a relatively small difference between the laser frequencies. This technique of optically generating microwaves by heterodyning is sometimes called “double beam modulation” since it involves the mixing of two optical beams to produce an output optical beam modulated at the beat frequency [22–24].

A variety of different techniques can be used to generate and mix two optical beams of the desired frequencies, and to convert the optical beat-frequency signal to a microwave signal. Two tunable DFB lasers can be monolithically integrated in close proximity to each other on the same chip, as shown in Fig. 21.5, and their beams allowed to mix on the surface of a photodiode.

The lasers shown in Fig. 21.5 can be electronically tuned by changing the current flowing in the side sections. Beat frequencies extending far into the mm-wave range can be generated in this manner. Alternatively, separate discrete tunable lasers can be used. For example Chau et al. [25] used light from two external-cavity tunable lasers emitting at approximately $1.55\text{ }\mu\text{m}$, combined by a 3-dB coupler and amplified by a high power EDFA, to generate microwaves with frequency as high as 95 GHz. A 1-nm optical band-pass filter was used to reduce the amplified spontaneous emission (ASE) noise from the EDFA. Of course, the photodetector that is used

to convert the optical beat frequency signal to microwaves must have an adequate bandpass to accommodate the frequency generated. In this case a velocity-matched distributed photodetector (VMDP) was used. The VMDP consisted of 4 MSM photodiodes, each with $14\text{ }\mu\text{m}$ length and $148\text{ }\mu\text{m}$ spacing between diodes. The VMDP exhibited a very high breakdown voltage, from 7 to 9 V, which corresponded to $3.5\text{--}4.5 \times 10^5\text{ V/cm}$. DC responsivity greater than 0.25 A/W at 5 V bias was measured. The response of the VMDP was relatively flat out to 105 GHz.

The lasers used in heterodyne microwave generation do not have to be of the DFB semiconductor type. Li et al. [26] have made a tunable mm-wave source by beating the output of two Nd:YVO₄/MgO:LiNbO₃ single mode microchip lasers. The devices consisted of a Nd:YVO₄ gain section and a MgO:LiNbO₃ tuning section monolithically integrated into a single composite crystal. The source can be tuned from DC to 100 GHz, with a sensitivity of 8.8 MHz/V and a tuning rate of over 10 THz/s.

The lightwaves out of two lasers do not necessarily have to be mixed externally in a photodiode to generate microwaves. Wang et al. [27] have demonstrated tunable mm-wave generation in two-section gain-coupled DFB lasers over a tuning range from 20 to 64 GHz, and microwave generation has also been reported in two-section index-coupled lasers [28]. The mechanism for generating microwaves in index-coupled lasers is thought to be nonlinear mode beating through interaction of exchanged photons with the carriers in the active media of both sections [28], while a large modulation index in gain-coupled lasers is considered to be related to enhanced transmission characteristics associated with the gain coupling [27]. In these two-section lasers the mixing occurs internally, so the output optical beam is modulated at the beat frequency. However, to generate microwaves some sort of optical-microwave converter, usually a junction or Schottky-barrier photodiode, must be used. The advantage of internal mixing is that optical injection locking creates a phase coherence between the beams that cancels phase noise.

Hsu et al. [29] have reported obtaining tunable dual-mode operation in a chirped grating DFB laser, which produces two gain-degenerate lasing modes in a single gain medium. Thus these modes can mix internally to yield an optical beat signal modulated in the microwave frequency range. A diagram of this dual mode laser is shown in Fig. 21.6.

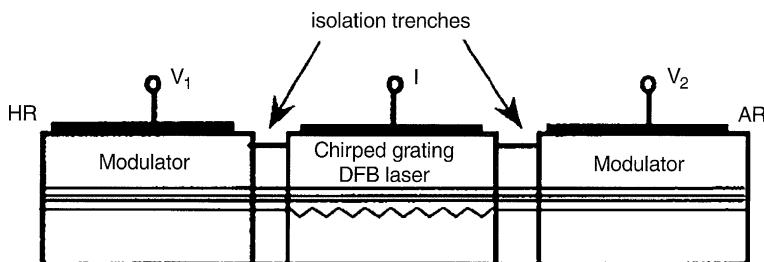


Fig. 21.6 Dual-mode chirped grating DFB laser [29] ©2000 IEEE

The device consists of a 300 μm long DFB laser section integrated with two 250 μm long modulator sections. Electrical isolation is provided by 80 μm long treched regions. One endface was high-reflectivity coated (HR) and the other was coated with an antireflecting (AR) coating. Selective biasing of the modulator sections enabled dual-mode operation of the laser to produce wavelengths of 0.5 nm and 1.4 nm, corresponding to frequency detunings of 63 GHz and 175 GHz.

A dual-mode laser has also been made by Grosskopf et al. [30]. In that case the device was composed of two DFB laser sections (each 250 μm long) with different Bragg wavelengths detuned by 2.5 nm, separated by a central integrated phase section that was 400 μm long. By varying the bias settings the output can be adjusted over a frequency range of 5 GHz to 50 GHz. This laser was used as the light source in an optical fiber radio-over-fiber transmission link in the 40 GHz band, over a 62 km path.

The problem of laser phase noise when the signals from two separate lasers are heterodyned can be mitigated by the use of a phase lock loop (PLL). An example of that approach is the fiber based optical injection phase locked loop (OIPLL) system described by Johansson and Seeds [31]. Their experimental arrangement is shown in Fig. 21.7.

The master laser (ML) is modulated at a third of the required millimeter-wave frequency. An adjustable delay line (AD) is used between the reference signal generator (RS) and the ML. Up to 20% of the ML output is injected into the slave laser (SL) diode by means of a splitter/coupler and an optical circulator (OC). The SL is tuned to lock to the third harmonic side-band of the injected ML light. Most of the light output from the circulator is combined with the remainder of the ML output in a second splitter/coupler to form the millimeter-wave modulated optical signal, which is fed to a spectrum analyzer and to a Fabry-Perot etalon. The remainder of the output from the circulator is coupled to a photodetector (PD) that detects the beat between the SL line and the residue of the ML fundamental line. The beat signal of the photodiode is amplified and the phase is compared to the reference signal in a

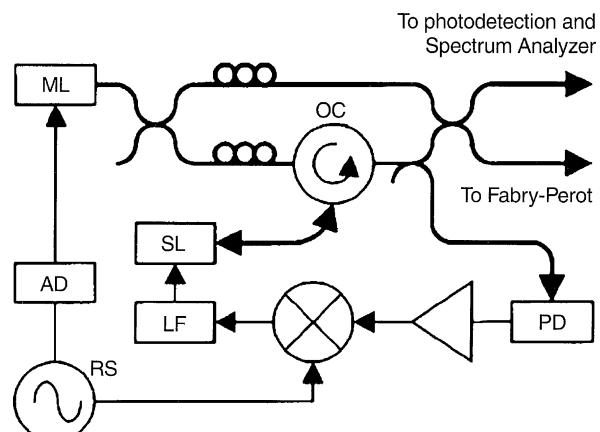


Fig. 21.7 Optical injection phase locked loop (OIPLL) system [31] ©2000 IEEE

subharmonically pumped double balanced mixer. The output of the mixer passes through a loop filter (LF) and is injected into the SL. This OIPLL system maintains a phase coherence between the ML and SL that allows phase noise cancellation. With two 1550-nm InGaAs-InGaAsP 16 quantum well DFB lasers used as ML and SL, a 36-GHz mm-wave signal wave generated with a linewidth of 220 MHz.

It is also possible to optically generate a microwave signal by heterodyning the output of a single laser after it has been split into two frequency-shifted components. The required splitting can be accomplished, for example, by using a standing wave surface acoustic wave optical modulator (SWSAWOM) such as the one discussed in Section 10.6 and shown in Fig. 10.9. Kissa et al. [32] have used that approach to implement microwave subcarrier multiplexing of optical signals onto the beam of a 816 nm wavelength diode laser. The integrated optic modulator was fabricated on a LiNbO₃ substrate, with a Ti-diffused planar waveguide. Three NTSC VHF TV signals (channels 2, 4 and 6) were upconverted by the modulator into the UHF range (channels 44, 47 and 49) to form a frequency division multiplexed (FDM) optical beam that was transmitted to a microwave spectrum analyzer for analysis and to a television receiver for viewing. The SWSAW optical modulator exhibited a Bragg diffraction efficiency of 41%. The transmitted spectrum consisted of a central component at twice the acoustic driving frequency of the modulator ($2f_a = 600$ MHz) and adjacent double side band (DSB) spectra, each composed of three groups of frequencies corresponding to the three channels.

Chen et al. [33] have demonstrated the generation of a mm-wave signal using only a single directly-modulated DFB laser in combination with a Mach Zehnder external modulator. Single-sideband operation was achieved by optical carrier suppression. The system was used to transmit a 2.5 Gb/s data channel over a 20 km path length.

21.4 Future Projections

The technical feasibility and economic viability of optical fiber interconnectors has clearly been demonstrated in a wide variety of applications. The use of optical fiber cable transmission systems is now widespread for interexchange links and undersea cables. Digital data transmission rates on the order of tens of gigabits per second are commonplace, and rates up to several terabits per second have already been introduced in commercial communications links and will become common within the next few years in DWDM systems. Optical fiber cables have not yet replaced copper wire cables in the subscriber loop in many locations but the plans for that conversion are already being made. When it occurs, the public telephone network will be transformed into a vastly more flexible medium for the delivery of all sorts of new services including interactive high-speed data transmission, high-definition TV cable channels, video telephone, high-speed FAX, and security and medical alarm systems. Many households are already enjoying the benefits of broadband data communication over the Internet and the World Wide Web via optical fibers.

With high-data-rate optical-fiber-cable transmission lines becoming well established, the need for optical integrated circuits has substantially grown. As a result, optical integrated circuits have ceased to be merely laboratory curiosities and have entered the marketplace. A number of companies are now marketing optical integrated circuits as off-the-shelf commercial products, and at least one company is offering to produce user-designed OICs in a “foundry” type of service such as has been available for many years for electronic IC chips. Presently the commercially available OICs are relatively limited — 2×2 optical switches, electro-optic phase modulators, Mach-Zehnder interferometer intensity modulators, multiplexers, demultiplexers, OADMs, couplers and splitters. However, these products have only been on the market for a relatively short time compared to electronic ICs. As systems engineers become more familiar with their use and their advantages, the demand for more sophisticated OICs will surely grow. As optical fibers are extended into the subscriber loops of telecommunications systems the demand for large quantities of high-speed optical switches and couplers can best be met by producing them in OIC form.

As fabrication technologies have improved and device dimensions have shrunk from micrometer to nanometer sizes, opportunities have become available for the development of more complex OICs within a smaller package. Many of these opto-electronic circuits will be fabricated within “photonic crystals”. This work is just beginning but a number of such circuits have already been realized. This new field of “nanophotonics” is described in Chapter 22.

The use of OICs and fibers as sensors is also a rapidly expanding field of application. So far most of the work has been confined to laboratory experiments and military applications, but a growing consumer market is evident. For example, the sensing and control functions in a modern automobile are becoming extremely complex and have been highly automated. Integrated optic sensors and signal processors with optical fiber interconnects can offer improved performance and reduced weight. One luxury car is already advertising a fiber optic instrumentation system.

Finally, the combination of integrated optic technology with microwave devices and systems, particularly in regard to wireless communications as discussed in Sections 21.1 and 21.2, has the potential of providing another expanding market, with applications in data transmission, wireless Internet, cellphones, avionics, security systems and television distribution.

Now that optical integrated circuits have emerged from the laboratory and some have become commercially available products which are being installed in systems “in the field”, a further impetus will be provided to continuing research and development of more sophisticated OICs.

Problems

- 21.1 A fiberoptic microwave link has the following characteristics. The overall transmitter efficiency is 10 mW/mA . The overall efficiency of the receiver is 3 mA/mW . The ratio of optical power @ transmit to the optical power @

- received is 2.5. The input resistance of the transmitter and the output resistance of the receiver are both 50Ω . What is the overall link gain?
- 21.2 If the 3 dB bandwidth of the fiberoptic microwave link of problem 21.1 is 40 GHz for a laser output power of 100 mW, what will it be if the laser power is increased to 200 mW?
- 21.3 Why are GaAlAs and GaInAsP materials well suited for making the devices that are components of OMMICS? What would you use as a substrate material with each of these?
- 21.4 If the current out of a fiberoptic microwave link is 30 times larger than the input current and the output resistance of the receiver is 1/2 the input resistance of the transmitter, what is the link gain in dB?
- 21.5 Name four major improvements in integrated optic technology that have occurred over the past 20 years that have led to modem lightwave telecommunications.
- 21.6 What are the advantages of using a substrate of GaAs or InP in an opto-microwave-monolithic-integrated-circuit (OMMIC)?
- 21.7 An analog microwave signal is transmitted over an optical fiber link by using the signal to modulate an optical beam and demodulating it at the receiver. The peak input current at the modulator is 10 mA and the peak output current at the demodulator is 15 mA. If the input resistance of the modulator and the output resistance of the demodulator are both 50Ω , what is the link gain in dB?
- 21.8 If the output resistance of the demodulator in Problem 21.7 is made to equal to 100Ω , what will be the new link gain in dB?

References

1. D. Politko, H Ogawa. The merging of photonic and microwave technologies. *microwave J.* March, 1992, 75 80
2. Y. Baeyens, A. Leven, W. Bronner, V. Hurm, R. Reuter, K. Kohler, J. Resenweig, M. Schleetweg: Millimeter-wave long-wavelength integrated optical receivers grown on GaAs. *IEEE photonics Tech. Lett* **11**, 868 (1999)
3. A. Rosen, P. Stabile, W. Janton, A. Gombar, P. Basile, J. Delmaster, R. Hurwitz: *IEEE Trans. MTT*-**37**, 1255 (1989)
4. A. Vaucher, W. Streiffer, C.H. Lee: *IEEE Trans. MTT*-**31**, 209 (1983)
5. P.R. Herczfeld, A.S. Durousch, A. Rosen, A.K. Sharma, V.M. Contarino: *IEEE Trans. MTT*-**34**, 1371(1986)
6. N.J. Gomes, A.J. Seeds: *Electron. Lett.* **23**, 1084 (1987)
7. R.G. Hunsperger, M.A. Mentzer: *SPIE Proc.* **993**, 204 (1988)
8. R.G. Hunsperger: *Proc. SBMO Int'l Microwave Symp*, Sao Paulo, Brazil (1989), IEEE Cat. No. 89TH0260-0, p. 743
9. R.W. Ridgway, D.T. Davis: *IEEE J. Lightwave Tech.* **4**, 1514 (1986)
10. IASTED/IEEE Int. Conf. On Wireless and Optical Communications (WC2002), Banff, Canada, July 17-19, 2002
11. P.R. Herczfeld: Applications of photonics to microwave devices and systems, in *Photonic Devices and Systems*, R.G. Hunsperger (ed). (Marcel Dekker, New York, 1994) Chap. 8

12. W.D. Jemison, P.R. Herczfeld, W. Rosen, A. Viera, A. Rosen, A. Paotella: Hybrid fiberoptic-millimeter-wave links. *IEEE Microwave Mag.* **1**, 44 (2000)
13. T. Kuri, K. Kitayama, A. Stohr, Y. Ogawa: Fiber-optic millimeter-wave downlink system using 60 GHz-band external modulation. *IEEE J. Lightwave Tech.* **17**, 799 (1999)
14. T. Kuri, K. Kitayama, Y. Takahashi: 60-GHz-band full-duplex radio-on-fiber system using two-RF-port electroabsorption transceiver. *IEEE Photonics Tech. Lett.* **12**, 419 (2000)
15. T. Olson: An RF and microwave fiber-optic design guide. *Microwave J.* August 1996, pp. 54–78
16. A. Umbach, T. Engel, H. -G. Bach, S. van Waasen, E. Droege, A. Strittmatter, W. Ebert, W. Passenberg, R. Steingrubler, W. Schlaak, G.G. Mekonnen, G. Unterborsch, D. Bimberg: Technology of InP-based 1. 55-μm ultrafast OEMMICS: 40-Gbit/s broad-band and 38/60-GHz narrow-band photoreceivers. *IEEE J. Quant. Electron.* **35**, 1024 (1999)
17. A. Stohr, K. Kitayama, T. Kuri: Fiber-length extension in an optical 60-GHz transmission system using an EA-modulator with negative chirp. *IEEE Photonics Tech. Lett.* **11**, 739 (1999)
18. K. Kitayama: Fading-free transport of 60 GHz-optical DSB signal in non dispersion shifted fiber using chirped fiber grating. *Proc. Int. Topical Meeting Microwave Photonics(MWP'98)*, Princeton, NJ, Oct. 1998, pp. 223–226
19. G. Smith, D. Novak: Broad-band millimeter-wave(38 GHz) fiber-wireless transmission system using electrical and optical SSB modulation to overcome dispersion effects. *IEEE Photonics Tech. Lett.* , **10**, 141(1998)
20. T. Kuri, K. Kitayama, Y. Takahashi:60-GHz-band full-duplex radio-on-fiber system using two-RF-port electroabsorption transceiver. *IEEE Photonics Tech. Lett.* **12**, 419 (2000)
21. R.A. Griffin, H.M. Salgado, P.M. Lane, J.J. O'Reilly: System capacity for millimeterwave radio-over-fiber distribution employing an optically supported PLL. *IEEE J. Lightwave Tech.* **17**, 2480 (1999)
22. C.S. Ih: All-optical communications and networks, in *Photonic Devices and Systems*. R.G. Hunsperger (ed) (Marcel Dekker, New York. 1994) Chap. 10
23. C.S. Ih: R.G. Hunsperger, J.J. Kramer, R. Tian, N. Wang, K. Kissia, J. Butlers A novel modulation system for optical communication. *Proc. SPIE* **876**, 30 (1988)
24. C.S. Ih: R.G. Hunsperger, X.L. Wang, J.J. Krarner, K. Kissia, R.S. Tian, J. Butler, X.C. Du, W.Y. Gu, D. Kopehik: Double beam modulation technologies II, (SPIEOE-LASE 90), Los Angles, CA. (1990) pp. 1218-53
25. T. Chau, N. Kaneda, T. Jung, A. Rollinger, S. Mathai, Y. Qian, T. Itoh, M.C. Wu, W.P. Shillue, J.M. Payne: Generation of millimeter waves by photomixing at 1. 55 μm using In GaAs-InAlAs-InP Velocity-matched distributed photodetectors. *IEEE Photonics Tech. Lett.* **12**, 1055 (2000)
26. Y. Li, A.J. C. Viera, P. Herczfeld, A. Rosen, W. Janton: Optical generation of rapidly tunable millimeter wave source. *Proc. International Topical Meeting on Microwave Photonics (MWP 2000)*, 11-13 Sept. 2000, pp. 259–262
27. X. Wang, W. Mao, M. Al-Mumin, S.A. Pappert, J. Hong, G. Li: Optical generation of microwave/millimeter wave signals using two-section gain-coupled DFB lasers. *IEEE Photonics Tech. Lett.* **11**, 1292 (1999)
28. H. Wenzel, U. Bandelow, H. Wunche, J. Rehberg: Mechanisms of fast self pulsations in two-section DFB lasers. *IEEE J. Quant. Electron.* **32**, 69 (1996)
29. A. Hsu, S.L. Chuang, T. Tanbun-Ek: Tunable dual-mode operation in a chirped grating distributed-feedback laser. *IEEE Photonics Tech. Lett.* **12**, 963 (2000)
30. G. Grosskopf, D. Rohde, R. Eggemann, S. Bauer, C. Bornholdt, M. Mohrle, B. Sartorius: Optical millimeter-wave generation and wireless data transmission using a dualmode laser. *IEEE Photonics Tech. Lett.* **12**, 1692 (2000)
31. L.A. Johansson, A.J. Seeds: Millimeter-wave modulated optical signal generation with high spectral purity and wide-locking bandwidth using a fiber-integrated optical injection phase-lock loop. *IEEE Photonics Tech. Lett.* **12**, 690 (2000)

32. K. Kiss, R.G. Hunsperger, C.S. Ih, X. Wang: Generation of microwave subcarriers for optical communication using standing-wave-surface-acoustic-wave waveguide modulator. Proc. Optical Society of America Annual Meeting, Oct. 30-Nov. 4, 1988
33. L. Chen, Y. Pi, H. Wen, S. Wen: All-optical mm-wave generation by using direct-modulation DFB laser and external modulator, *Microwave Opt. Technol. Lett.* **49**, 1265 (2007)

Chapter 22

Nanophotonics

The field of nanotechnology, which deals with particles and structures having dimensions on the order of nanometers (10^{-9} m), is very broad. It includes work done in the fields of chemistry, physics, biology, materials science, mechanical engineering, electrical engineering, environmental science and even ethics. However, in keeping with the focus of this book, the material to be covered in this chapter will be confined to topics relevant to integrated optics.

22.1 Dimensions

In the previous chapters of this book the topics covered might be described as belonging to the field of microphotonics, since they involved photons of light interacting with physical structures having dimensions mostly on the order of micrometers. The two exceptions were gratings and quantum wells, some of which had periodicity on the order of 100 nm. As fabrication techniques have progressed it has become possible to produce nanometer-sized structures such as quantum wires, quantum dots, holographic optical elements (HOEs) and photonic crystals (PhCs). This chapter gives a review of the basic theory and technology of this new field of nanophotonics. The introduction of sub-micron-sized structures into photonic devices and integrated circuits has produced improvements in performance, as noted in Table 22.1.

Ultimately, photonic integrated circuits (PICs) fabricated in PhCs may be the long-desired solution to the “electronic bottleneck” that currently is limiting the performance of personal computers to data rates of a few GHz. By transmitting signals with photons rather than electrons in PhCs, it might be possible to build a computer that operates at terahertz (10^{12} Hz) data rates [1].

22.2 Properties of Electrons and Photons

In Chapter 11 the particle-like properties of photons and electrons were used to describe a microscopic model of light generation and absorption in materials.

Table 22.1 Benefits of nanostructures

-
- *Improved performance of lasers*
 - Higher quantum efficiency
 - Narrower linewidth
 - Greater range of wavelengths
 - Improved thermal stability
 - *Improved performance of detectors*
 - Higher quantum efficiency
 - Greater bandwidth
 - *Improved performance of modulators*
 - Higher modulation index
 - *Greater device density on chip*
-

When the size of features in the material within which the photons and electrons are interacting approach nanometer dimensions, their wave-like properties become much more important. Both transmission and confinement of electrons and photons depend on their wave-like properties of frequency, wavelength and phase. The wavelengths of the photons that are commonly used in PICs are in the range of 0.6–1.8 μm , while the wavelengths of electrons in a crystal are on the order of Angstroms.

In free space, both electrons and photons can be described by a propagating plane wave with constant amplitude A and a propagation vector \mathbf{k} which gives the direction of propagation, as shown below

$$\Psi = A(e^{i\mathbf{k}\cdot\mathbf{r}-\omega t} + e^{-i\mathbf{k}\cdot\mathbf{r}+\omega t}), \quad (22.1)$$

where \mathbf{r} is the radius vector, ω is the radian frequency and t is time. The wave vector \mathbf{k} is related to the momentum \mathbf{p} , as discussed in Chapter 11, by

$$\mathbf{p} = \hbar\mathbf{k} = \hbar\mathbf{k}/2\pi, \quad (22.2)$$

where \hbar is Planck's constant. The magnitude of the wave vector \mathbf{k} is given by

$$|\mathbf{k}| = 2\pi/\lambda = \omega/v, \quad (22.3)$$

where v is the velocity of propagation.

In a solid material, the above three free-space equations are basically still valid for both photons and electrons but the amplitude and wavelength are modified by interactions with the atoms and ions of the solid. This results in a change in the velocity of propagation, which is characterized by the index of refraction for photons and the electrical resistance to the flow of electrons. In crystalline materials, either semiconducting or dielectric, there are strong scattering and resonance mechanisms that can result in localization of electrons and photons and in blocking of propagation for certain unallowed values of \mathbf{k} .

Despite the many similarities between the behavior of photons and that of electrons, there are some differences. The wavefunctions of photons correspond to electromagnetic vector fields \mathbf{E} and \mathbf{H} , while the wavefunctions of electrons correspond to scalar fields representing the probability that an electron is at a certain location with a certain energy. Electrons have electrical charge and a rest mass, while photons do not. Electrons have characteristic spin states, while photons have polarization.

22.3 Confinement of Photons and Electrons

In Chapter 2 the confinement of photons to a waveguide was described. The waveguide consisted of a region of higher index of refraction surrounded by regions of lower index of refraction. This confinement resulted in a distribution of optical energy in one or more dimensions that remained constant over time, i.e., an optical mode or set of optical modes. Modes of a one-dimensional planar waveguide are shown in Fig. 22.1.

The field distributions and the propagation constant for each mode are obtained by the solution of Maxwell's equations and the matching of boundary conditions as explained in Chapter 3. Notice that there are evanescent "tails" to the modes that extend into the regions of lower index of refraction. This phenomenon can be described as photonic tunneling, since some photons of the mode are penetrating the barrier created by index of refraction difference. The greater the index difference at the interface, the fewer the number of photons that can tunnel through the barrier.

In Chapter 18 the confinement of electrons to a quantum well was described. The confinement of an electron to a region of lower potential energy in a material of narrower bandgap surrounded by a region of wider bandgap material results in a quantization of allowed electron energies into discrete energy levels n_1, n_2, n_3, \dots . The diagram of the wavefunctions Ψ of electrons confined in a quantum well is very similar to that of photons confined in a waveguide, as shown in Fig. 22.2.

In this case, the evanescent tails of the wavefunctions represent the probability that an electron will be found at a point within the barrier region. Since the tails are nonzero there is a probability that the electron can tunnel out of the well a certain distance into the barrier material. The probability density P is a function of position, given by

Fig. 22.1 Diagram of the confined modes of a planar waveguide with $n_2 > n_3 > n_1$. The distributions of the electric field in the x -direction for three modes are shown. The z -axis represents the direction of propagation

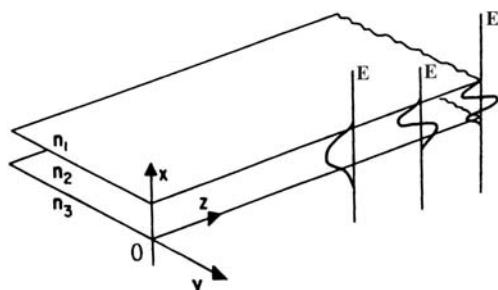
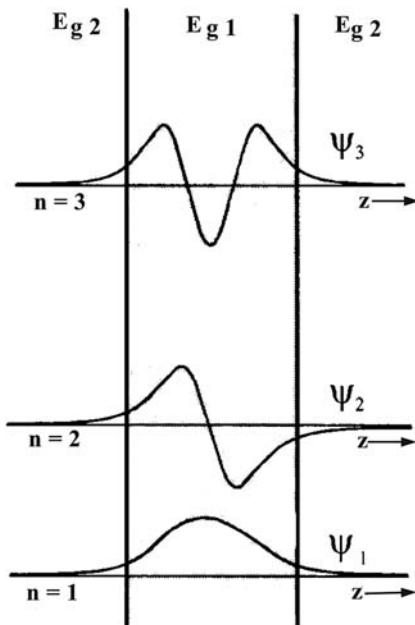


Fig. 22.2 Diagram of the wavefunctions Ψ of an electron in the quantized energy levels n_1 , n_2 and n_3 of a quantum well formed by a material of bandgap E_{g1} and barrier regions of bandgap E_{g2} , where $E_{g2} > E_{g1}$.



$$P(z) = |\Psi(z)_n|^2. \quad (22.4)$$

The wavefunctions Ψ_n and the energy levels $E_{n,n}$ are obtained by the solution of Schrödinger's equation [2] as described in Chapter 18. It is interesting to note that the confinement of both electrons and photons is best described by considering their wave nature rather than their particle nature.

The preceding paragraphs have described one-dimensional confinement of electrons and photons. However, confinement can be extended to two and three dimensions as well. Electrons can be confined to two-dimensional quantum wires [3] and three-dimensional quantum dots [4]. Similarly, a channel waveguide or an optical fiber waveguide produces two-dimensional confinement of photons, and a microsphere optical cavity [5] yields three-dimensional confinement. These singular confining structures can be combined to produce a multiplicity of coupled, confining structures, such as multiple quantum wells (Chapter 18) and quantum dot arrays [6].

22.4 Photonic Crystals

22.4.1 Classes of Photonic Crystals

When photon confining structures are multiply coupled on a single chip so as to produce a periodically varying index of refraction (dielectric constant) with a

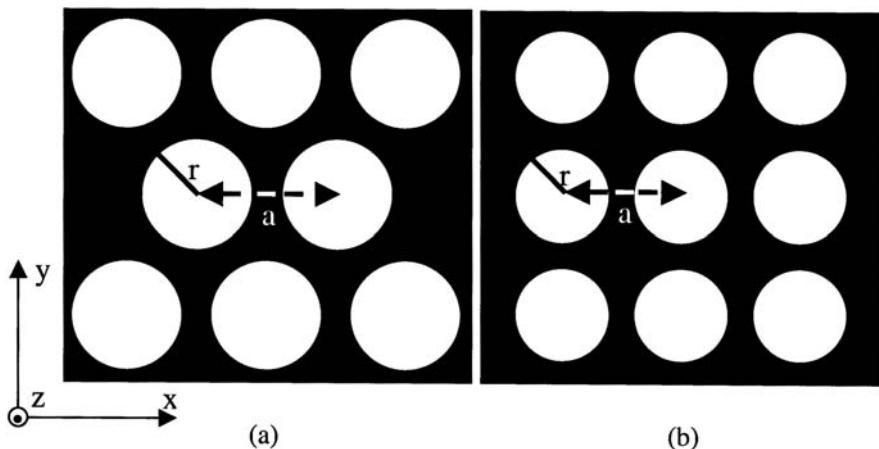


Fig. 22.3 (a, b) Diagram of common two-dimensional photonic crystal structures: (a) a triangular lattice and (b) a square lattice [7]

periodicity on the order of the wavelength of the photons, a photonic crystal is formed. An example of photonic crystal lattices is shown in Fig. 22.3. For light of $1\text{ }\mu\text{m}$ vacuum wavelength, both the lattice constant a and the radius of the perturbation r would typically be several hundreds of nanometers.

Photonic crystals can be either one-, two- or three-dimensional. Technically speaking, examples of one-dimensional photonic crystals are a diffraction grating like that shown in Fig. 15.10 and a multilayer dielectric mirror like that of Fig. 15.14. However, since grating structures like those have been used for many years before the advent of modern photonic crystals, it is not common to call them photonic crystals.

Two-dimensional photonic crystals are usually fabricated by producing a lattice of air holes (lower index of refraction) in a layer of higher-index material or by producing a lattice of higher-index rods in a layer of lower-index material (such as air). Lattices of these two types are shown in Fig. 22.4. These lattice structures require a very large aspect ratio. (The aspect ratio of a feature on an integrated circuit chip is defined as its height divided by its width.) However, such high-aspect-ratio structures can be fabricated fairly routinely by using the recent advances in nanolithography and e-beam and ion-beam sputter etching that are described in Section 22.5.4.

22.4.2 Comparison of Electrons in Semiconductor Crystals to Photons in Photonic Crystals

The name, photonic crystal, derives from the fact that the behavior of photons in a photonic crystal is very similar to that of electrons in a crystalline semiconductor.

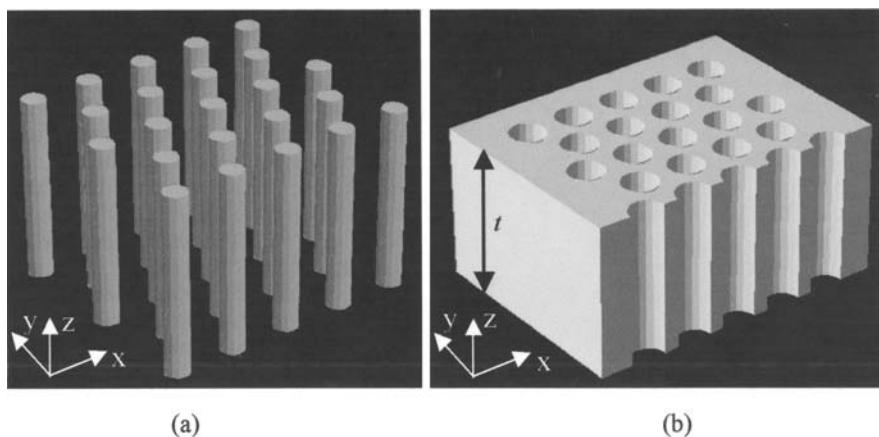


Fig. 22.4 (a, b) Two-dimensional photonic crystals: (a) dielectric rods in air and (b) holes in higher-index material [7]

In moving through a semiconductor crystal, an electron is affected by the periodic potential produced by its Coulomb interaction with the atoms of the semiconductor. Solution of the Schrödinger equation for the wavefunction of the electron and its allowed energy states yields the familiar dispersion diagrams of allowed electron energy states vs. wave vector, such as those shown in Fig. 22.5.

Key observable features are the valence and conduction bands and the bandgap E_g , within which there are no allowed electron energy states. When Maxwell's equations are solved for the case of photons in a photonic crystal, very similar dispersion diagrams of energy (frequency) vs. wave vector result, including the feature of a "bandgap" of frequencies within which no allowed states exist, i.e., no propagation can occur. The solution of Maxwell's equations for photons in a periodic dielectric lattice is a fairly complicated process. There are no simple closed-form equations that can be easily derived. However, accurate solutions can be obtained using computer-aided methodologies, such as the plane wave method (PWM) [8, 9] and finite-difference time-domain (FDTD) method [10]. These mathematical methods can be adapted to include the vector nature of electromagnetic fields as they

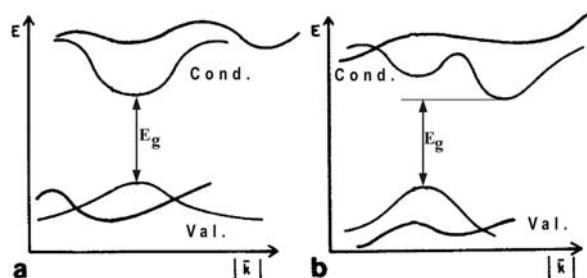
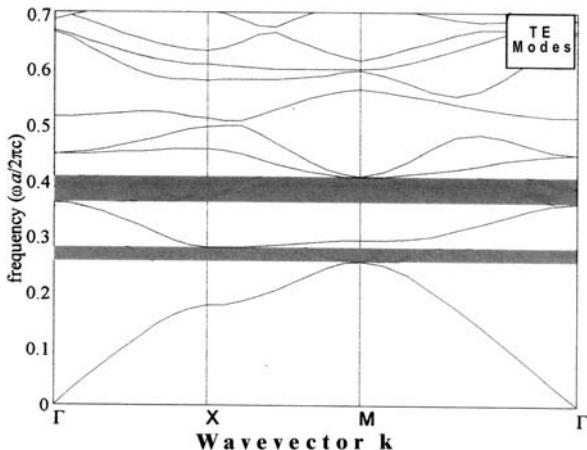


Fig. 22.5 (a, b) Electron energy E vs. wave vector \mathbf{k} in a semiconductor: (a) in a direct bandgap material and (b) in an indirect bandgap

Fig. 22.6 Dispersion diagram for TE polarized photons of radian frequency ω propagating in a rectangular dielectric lattice with lattice constant a



propagate within periodic lattice structures. The PWM works well for structures that have uniform periodicity and can be used for some structures that contain defects, as long as the defects can be incorporated into a periodic “super cell” [11]. However, structures containing non-periodic sets of defects, such as the waveguides and splitters discussed in Sections 22.7.1 and 27.7.2 are best handled by the FDTD method, which does not require periodicity. A typical dispersion diagram calculated for photons propagating in a rectangular dielectric lattice is shown in Figure 22.6.

The corner points of the irreducible Brillouin zone in k -space Γ ($=0$), X ($=\pi/x$) and M ($=\pi/(x+y)$) are marked in the diagram. (For a review of the Brillouin zones of a crystal, see, for example, Ibach and Lüth [12] or Kittel [13].) The gray bands mark the bandgaps for TE photons, within which there are no allowed states, i.e., propagation of photons of those frequencies is blocked. Figure 22.6 is a plot of allowed propagation energies (frequencies) of photons in a PhC vs. their wave vectors, just as Fig. 22.5 is a plot of the allowed energies of electrons in a semiconductor crystal vs. their wave vectors.

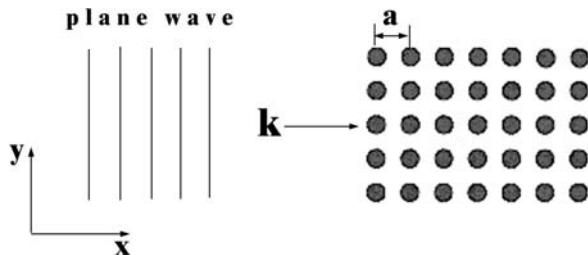
The blocking of the propagation of photons of certain frequencies in a photonic crystal can be easily understood without going through the full calculation of the dispersion diagram by considering it to be a phenomenon of Bragg scattering (reflection) similar to the reflection of photons in a DBR laser, as discussed in Chapter 15. Consider a plane wave traveling in two-dimensional rectangular lattice as shown in Fig. 22.7.

The \mathbf{E} and \mathbf{H} fields of the wave are given by

$$\mathbf{E}, \mathbf{H} \sim e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \text{ where} \quad (22.5)$$

$$|\mathbf{k}| = \omega/v = 2\pi/\lambda. \quad (22.6)$$

Fig. 22.7 Diagram of plane wave traveling in a rectangular dielectric lattice



Remember, the basic condition for Bragg scattering through an angle $2\theta_B$ is given by

$$\sin\theta_B = \lambda/2\Lambda, \quad (22.7)$$

where θ_B is the Bragg angle and Λ is the periodicity. In this case $\Lambda = a$; thus waves of wavelength equal to $2a$ will produce $\sin\theta_B = 1$ and will be scattered through an angle $2\theta_B = 180^\circ$. They will be reflected and will not propagate.

In a real PhC, small perturbations will spread the condition $\lambda = 2a$ over a narrow range of wavelengths approximating $2a$ to yield a bandgap. The transmission spectrum of a silicon PhC is shown in Fig. 22.8 [14]. It consists of a triangular lattice of air holes in a silicon layer with $a = 437$ nm and $r = 175$ nm. The resulting bandgap, with edges defined by transmission reduced by 20 dB, ranges from 1118 to 1881 nm. There is only 1% transmission (or less) permitted within this range of wavelengths.

For a two-dimensional PhC like this one, a bandgap may or may not exist. Sometimes the ω vs. k curves overlap so that there is no bandgap; it depends on the particular values of spacings and dielectric properties. A large r/a ratio and large

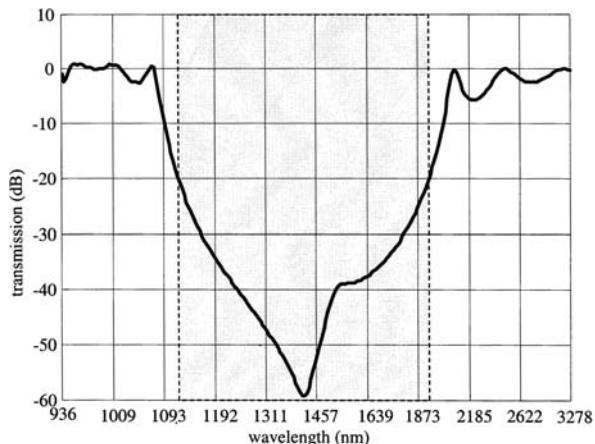
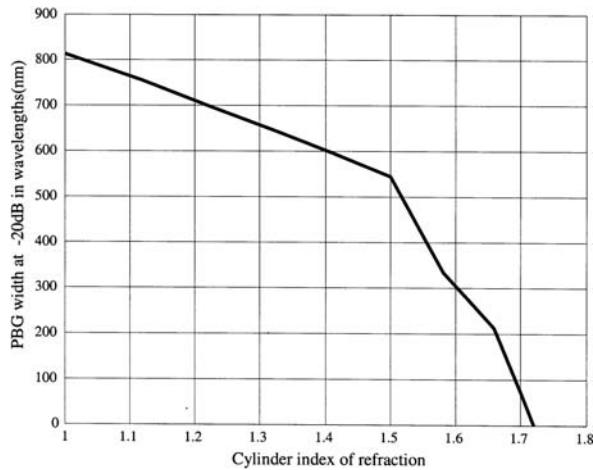


Fig. 22.8 Transmission spectrum of a silicon photonic crystal [14]

Fig. 22.9. Reduction of bandgap as a function of cylinder index of refraction [14]



difference in permittivities (dielectric constants) are required. (A one-dimensional PhC like a grating or a multilayer dielectric mirror always has a bandgap.) When the air holes of the PhC of Fig. 22.8 were filled with a material with dielectric constant > 1 , the width of the bandgap was reduced, as shown in Fig. 22.9 [14].

Since the index of refraction of silicon is about 3.5 for the wavelengths used in PICs, the data of Fig. 22.9 demonstrate that a large index difference is required in order to produce a bandgap. When defects are controllably introduced into PhCs, disrupting the regularly ordered lattice, new propagation states are created that can be used to create devices such as waveguides, couplers, switches and filters. Such devices are discussed in Section 22.7, but first it is important to review the fabrication technologies that are used to process nanostructures.

22.5 Fabrication of Nanostructures

There is a great variety of methods that have been demonstrated for producing either structures with features having dimensions on the order of nm or particles of nm size. For example, plasma arcng, sol-gels, electro-deposition, chemical vapor deposition, ball milling, self-assembly and natural nanoparticles have all been used. See, for example, Wilson et al. [15]. However, these have not found much application in the field of photonics. Instead, the traditional methods of semiconductor device and PIC fabrication, such as molecular beam epitaxy (MBE), metalorganic vapor phase epitaxy (MOVPE), e-beam lithography, ion-beam etching and reactive-ion etching (RIE), have been upgraded and adapted to the fabrication of nanophotonics structures and PhCs.

22.5.1 Molecular Beam Epitaxy

The basic principles and technology of MBE are discussed in Section 4.4.4, and the use of MBE to create multiple-layer, multiple quantum well (MQW) structures in a number of photonic devices is described in Chapter 18. Certain additions and modifications must be made to the basic MBE system shown in Fig. 4.13 to qualify it for nanolayer fabrication. The remotely operated shutters covering the heated effusion-cell sources must be capable of switching open and closed in 0.1 s, or less, in order to produce layers of thickness < 100 nm. The substrate must be rotated during growth to insure uniformity of the grown layer. The temperature of the substrate must be controlled to 0.1° accuracy to stabilize the growth rate. Also, the system should be fitted with analysis tools such as a mass-spectrometric flux monitor to measure layer composition and reflection high-energy electron diffraction (RHEED) to measure layer crystallinity. These measurements must be made while growth is going on in order to control the quality of the layer.

The layers grown by MBE can be patterned laterally to produce two-dimensional devices such as quantum wires, quantum dots or planar PhCs. This can be done by either growing on pre-patterned substrates or patterning the layer after growth by lithographic masking, exposure and etching.

MBE growth can be used for III–V and II–VI semiconductors as well as for group IV semiconductors. Thin layers of both ternary and quarternary compounds, such as $\text{Ga}_x\text{Al}_{1-x}\text{As}$ and $\text{In}_{1-y}\text{Ga}_y\text{As}_{1-y}\text{P}_y$, can be grown by MBE.

22.5.2 Metalorganic Vapor Phase Epitaxy

In MOVPE the constituent atoms are delivered to the substrate as a gaseous flow within a growth reactor furnace. A basic horizontal-flow MOVPE reactor is shown in Fig. 22.10.

MOVPE is a form of metalorganic chemical vapor deposition (MOCVD) in which the growth is epitaxial. The metalorganic gasses such as triethyl indium (TEIn) and triethyl gallium (TEGa) are usually transported along with H_2 as a carrier gas [16]. A crystalline substrate must be provided to act as the seed for epitaxial growth. MOVPE is used for the growth of thin layers of III–V and II–VI semiconductors in order to maintain a high enough vapor pressure of the constituent elements to prevent dissociation of the layer during growth. It is not used for silicon vapor phase epitaxy because it is not necessary. Silicon is an elemental semiconductor which can be grown by conventional VPE using silicon tetrachloride, or silane, dichlorosilane and trichlorosilane as source gases.

For the growth of nanolayers by MOVPE special efforts must be made to provide the uniformity and control of growth rate that is required. Temperature within the reactor must be stabilized to within 0.1° and the remotely controlled valves must be switchable within 0.1 s. A premixing chamber must be added at the reactor input to insure uniformity of the layer.

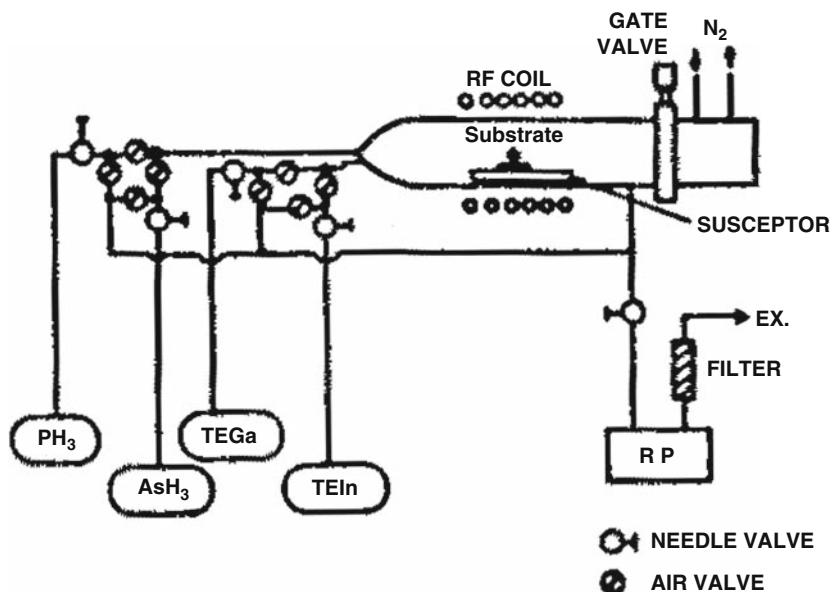


Fig. 22.10 A schematic diagram of a basic horizontal-flow MOVPE reactor

Despite the apparent simplicity of the basic system shown in Fig. 22.10, practical MOVPE systems are quite complex and are usually purchased commercially rather than being assembled by the researcher in the laboratory. The major reason for this complexity is that the source gases are extremely dangerous and must be handled with strict adherence to established standards. In addition to a number of horizontal-flow MOVPE reactors that are marketed commercially, a closed coupled showerhead reactor is available. (The term Close Coupled Showerhead® is a trademark of Aixtron.) The showerhead consists of an array of uniformly distributed injection holes, spaced 1–2 cm away from the substrate surface. This type of reactor has been shown to be less sensitive to growth conditions like reactor temperature, pressure and rotation [17].

22.5.3 Nanoscale Lithography

In regard to semiconductor device and PIC fabrication, lithography is the technology that is used to transfer the lateral dimensions of structures to the active layer. The basic approach used by the semiconductor industry is to coat the layer with a “resist” that is sensitive to either photons of light or electrons such that it becomes either removable in a developer solution when exposed (positive resist) or nonremovable (negative resist). The resist is exposed through a mask that defines lateral structures by blocking some areas and passing light through to others. The resist is “developed” in a solution that dissolves removable areas. The developed resist is

then used as a mask against either liquid or sputter etching. The size of the features that can be created depends strongly on the wavelength of the exposure source. As one might expect, shorter wavelengths yield smaller feature sizes. The diffraction-limited resolution of a circular lens is given by the empirically determined Rayleigh criterion, which states that

$$\sin\theta = 1.220\lambda/D, \quad (22.8)$$

where θ is the angular resolution, λ is the vacuum wavelength and D is the diameter of the lens. This leads to an expression for the spatial resolution Δl of an ideal lens, given by

$$\Delta l = 1.220 f\lambda/D, \quad (22.9)$$

where Δl is the size of smallest object that the lens can resolve and f is the focal length of the lens. Equations (22.8) and (22.9) assume an ideal lens. In a practical lithographic projection exposure system, imperfections in the optical train and the resist further limit the resolution. This results in the following adjusted form of the Rayleigh criterion [18]:

$$W_{\min} \approx k f\lambda/D \approx k \lambda/NA, \quad (22.10)$$

where W_{\min} is the minimum feature size that can be produced, NA is the numerical aperture of the optical system and k is a constant that characterizes the ability of the resist to distinguish between small changes in intensity. Typically k is approximately 0.75. The effect of wavelength on minimum feature size W_{\min} based on Eq. (22.10), assuming $k = 0.75$ and $NA = 0.6$, is shown in Table 22.2

For x-ray and e-beam exposure the assumed values for k and NA are probably optimistic. Empirical data indicate a currently achievable $W_{\min} = 10$ nm for both of these methods.

From the data in Table 22.2 it can be seen that ultraviolet (UV) and deep ultraviolet (DUV) exposure systems are not suitable for fabricating nanoscale features. Extreme ultraviolet (EUV) systems are a possibility. However, because of high absorption associated with the very short wavelength no lenses are available. All focusing and directing must be done with mirrors and must be done in a vacuum to

Table 22.2 The effect of exposure wavelength on minimum feature size

Exposure source – Wavelength (nm)	W_{\min} (nm)
UV Hg Arc i-line – (365)	456
DUV F ₂ laser – (157)	196
EUV laser plasma – (13.4)	16.8
X-ray – (0.8)	1.0
e-Beam – (0.07)	0.09

avoid absorption and scattering of the photons. For a review of extreme ultraviolet lithography (EUVL), see, for example, Bjorkholm [19].

X-ray lithography can be used for nanostructure fabrication. In that case, special materials must be used in the exposure mask. Materials that are transparent to x-rays, and hence can be used as mask substrates, are silicon, silicon carbide, silicon nitride, boron nitride and diamond. Materials that are strong absorbers of x-rays, and hence can be used to make blocking regions of the mask, are gold and tungsten. X-ray exposure systems are complicated by the fact that there are no known effective lenses or mirrors for x-rays. Thus the beam cannot be collimated and is radiating outward from a small point source to the larger mask. This introduces errors in the shape and position of the features that complicates mask design [18]. A collimated x-ray beam that avoids these problems can be obtained from a synchrotron source, but they are very large and expensive.

The most widely used method for patterning the lateral dimensions of nanoscale photonic devices has been electron-beam (e-beam) lithography. Electron beams consist of charged particles for which the methods of collimating, focusing and directing have been well developed. There are also well-known electron-resist materials such as polymethyl methacrylate (PMMA) and a number of proprietary commercial resists. The use of an e-beam for exposure of nanoscale features is consistent with the use of short-wavelength photons. An electron traveling at a velocity of $v = 10^7$ m/s has a De Broglie wavelength of

$$\lambda = h/mv = 6.63 \times 10^{-34} / (10^7 \times 9.11 \times 10^{-31}) = 7.2 \times 10^{-11} \text{ m}, \quad (22.11)$$

or 0.07 nm.

E-beam exposure is most often performed by direct writing using a focused beam with its lateral position relative to the target being controlled by a preprogrammed computer. This approach avoids the need for a mask and eliminates problems associated with masks, such as distortion and wear. A basic e-beam exposure system is shown in Fig. 22.11.

The metal film is necessary to provide a grounding path for electrons. Otherwise a charge buildup on the surface of the resist would deflect the beam. Focused beam diameter can be as small as 5 nm, with minimum linewidth = 20 nm. Accelerating voltage is typically 25–50 kV. Direct-write e-beam lithography is very effective for use in the laboratory situation, where device design is frequently changed and volume production is not an issue. However, it is a very time-consuming operation as the beam is stepped from point to point by the computer. When PhCs and other nanoscale photonic devices go into large-volume production, it is likely that EUVL will be used when possible.

22.5.4 Nanomachining

Once the lateral dimensions of the nanostructure have been defined by lithography, it is necessary to controllably remove material in certain regions to shape the

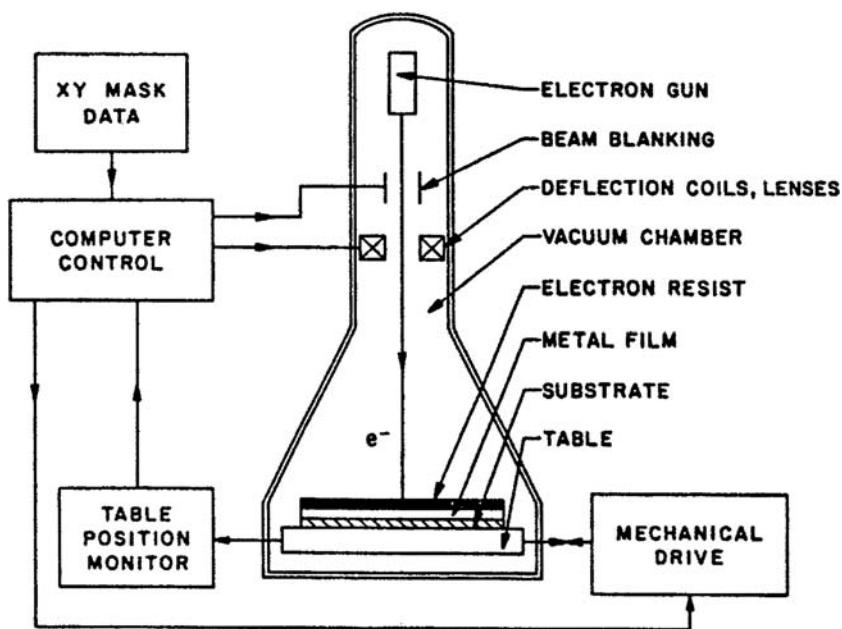


Fig. 22.11 Diagram of a basic e-beam exposure system

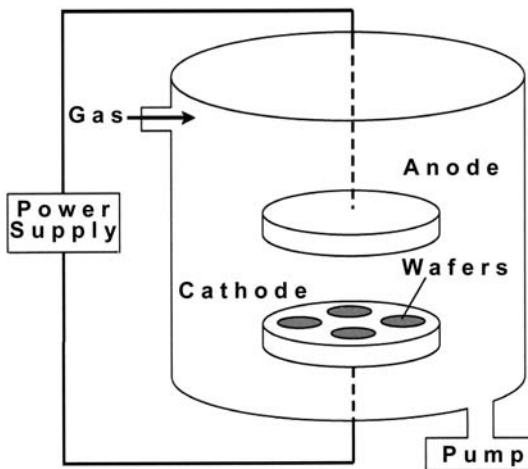
nanostructure. This might involve making a lattice of holes in a layer as in Fig. 22.4b or shaping columns as in Fig. 22.4a. Because of the high aspect ratios and close spacing desired in most cases, liquid etching is inadequate and special etching techniques must be used.

The basic method used is sputter etching, sometimes called dry etching. In sputter etching the layer to be etched is bombarded with energetic ions. Usually the ions of inert gases such as argon or krypton are used. Ion energy is typically 1–2 keV. The impinging ions transfer energy to the atoms near the surface of the layer causing them to be ejected. The rate of removal depends on the ion flux (number of ions/area·s) and the sputtering yield, which is defined as the number of atoms ejected from a target surface per incident ion. Sputtering yield generally increases with increasing ion energy up to about 2 keV, then decreases due to ion implantation. Heavier “bullet” ions generally increase sputtering yield for the same energy. Measured values of sputtering yield for a variety of the commonly used ions and target materials have been published, see, for example, Vossen and Kern [20]. There are two major types of sputtering that are used, plasma and ion beam.

In plasma sputtering the ions are generated by establishing an electrical “glow” discharge between an anode and a cathode within a vacuum system containing one of the inert gases. A diagram of a plasma sputter etch system is shown in Fig. 22.12.

The pressure is typically $(2\text{--}20) \times 10^{-3}$ torr. The high voltage applied between the anode and the cathode can be either dc or rf with a dc component. A dc negative bias must be maintained on the cathode so that ions will be attracted to it. The

Fig. 22.12 Diagram of a plasma sputter etching system



impinging ions sputter away atoms from the layer in regions where it is not masked by the developed resist. The ions do cause some surface damage and also some ions may become implanted, particularly at higher sputtering energies. However, if these effects prove detrimental to device performance they can be mitigated by annealing at about 200°C for 10 or 20 min.

An improvement can be made to the basic plasma sputter etching system of Fig. 22.12 by adding magnets so as to produce a magnetic field with its field lines aligned perpendicularly to the anode and cathode. This field causes the electrons in the plasma to travel a spiraling path, thus increasing the probability that they will encounter a gas atom and ionize it. (The path of the ions is not strongly affected by the magnetic field because of their much greater mass.) This type of sputtering is called magnetron sputtering. It results in the generation of a high-density plasma and increases the sputtering rate. It also allows reduction of the chamber pressure to the 10^{-5} – 10^{-3} torr range, which reduces stray deposition of sputtered atoms within the system.

The second major type of sputter etching is ion-beam sputtering. In ion-beam sputtering the ions are generated in a localized ion source, extracted and then accelerated and focused to form a collimated beam as shown in Fig. 22.13. The inert gas ions are generally used just as in plasma sputtering.

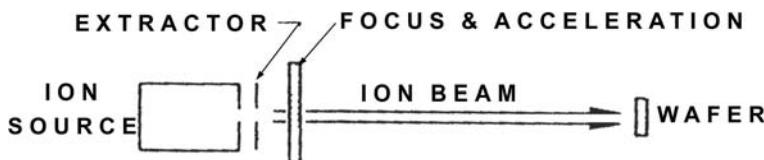


Fig. 22.13 Diagram of ion-beam sputter etching system

One advantage of ion-beam etching as compared to plasma etching is that the beam is well collimated. Therefore there is less undercutting and higher aspect ratios can be achieved. Also, the ion-beam system can (and should) be operated in a hard vacuum of 10^{-7} or better. This reduces contamination. There is no incorporation of stray-contaminant atoms that might be present in the gas of a plasma sputtering system. The ions extracted from the source can be mass filtered for purity as in an ion implantation system. Another advantage of ion-beam sputtering is that the ion energy and beam current can be independently controlled by varying the accelerating voltage and extractor voltage, respectively. For these reasons ion-beam sputtering is generally preferred for fabrication nanostructures.

Sometimes chemically active ions are used rather than inert gas ions in both plasma and ion-beam sputter etching. This is called reactive-ion etching (RIE). Chemically active ions can be conveniently generated in a plasma of halogen or halide gases such as SiCl_4 , Cl_2 , CHF_3 , CF_4 , C_2F_6 and SF_6 . Gas pressure is typically maintained in a range between 10^{-3} and 10^{-1} torr by adjusting gas flow rates. RIE is a particularly anisotropic etch technique capable of producing aspect ratios of more than 20:1 in some processes. These are called deep reactive-ion etch (DRIE) processes. There are two commonly used DRIE processes, cryo-DRIE [21, 22] and Bosch DRIE [23, 24].

In cryo-DRIE, the wafer is chilled to -110°C (163 K). The low temperature slows down the chemical reaction that tends to produce isotropic etching of trench sidewalls. However, ions continue to bombard the bottom surface of the trench and etch it away. This process produces trenches with vertical sidewalls.

The Bosch process, patented by Bosch GmbH [25], features multiple cycles of alternating etches and depositions. The trenches are formed in small increments until the desired aspect ratio is reached. The deposition protects the sidewalls during each etch step, preventing any lateral etching that would reduce the anisotropy of the process. A diagram of a modified Bosch process used at the University of Delaware is shown in Fig. 22.14 [7, 26].

Both cryo-DRIE and Bosch Drie are capable of producing high-aspect-ratio structures with vertical side walls.

A modified plasma sputtering technique that is also capable of producing a high aspect ratio and vertical side walls is inductively coupled plasma (ICP) etching. ICP etching is a technique which uses rf energy coupled into a low-pressure gas such as SF_6 by an inductive coil mounted on the outside of a non-conducting, insulating window. There are two types of ICP geometries: planar and cylindrical. In planar geometry, the electrode is a coil of flat metal wound like a spiral. In cylindrical geometry, it is like a helical spring. When an rf current is passed through the coil, the resulting time-varying magnetic field induces breakdown of the gas within the reactor chamber and creates a high-density plasma. Two separate power supplies can be used in an ICP system for plasma sputter etching. This gives independent control of the ion energy and the ion-beam current. A low-frequency supply is used to generate a high-density plasma and a high-frequency supply is used to provide a wafer bias. This dual power supply arrangement allows the generation of a high-density plasma for increased etch rate, along with a low wafer bias for low damage to

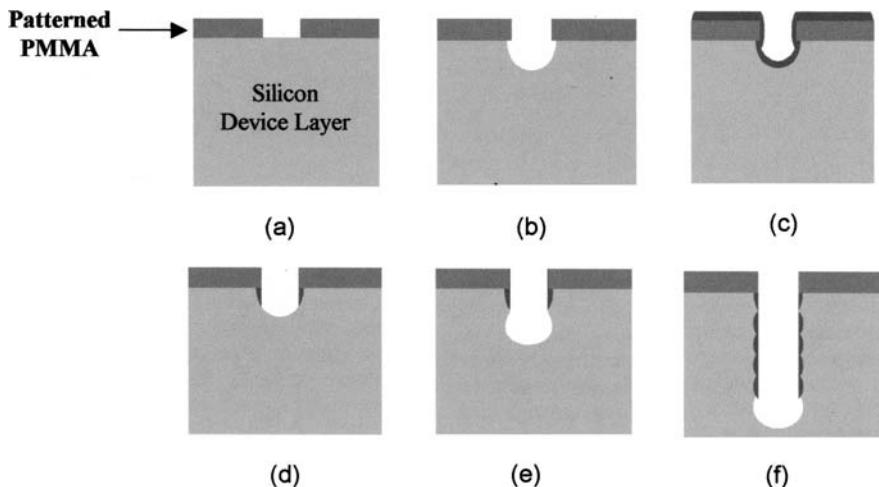


Fig. 22.14 Diagram of a modified Bosch etch process [26]: (a) pattern and develop PMMA resist layer; (b) high-pressure, low-energy isotropic etch; (c) deposition of CHF_3 passivation layer; (d) high-energy, low-pressure anisotropic etch via ion bombardment; (e) repeat the high-pressure, low-energy isotropic etch and cycle through steps (b–d); and (f) etch profile after multiple repetitions of steps (b–d)

the wafer surface. The wafer bias can be adjusted to control the decree of anisotropy of the etch process. In a conventional capacitively coupled sputter etch system with a single power supply, both the ion energy and the beam current are determined by the single applied voltage.

22.6 Characterization and Evaluation of Nanostructures

22.6.1 Available Tools

There is a wide selection of analysis tools available for the characterization of nano-materials, nanostructures and nanodevices. These are the standard instruments that are found in many laboratories, so a detailed description of all of them is not necessary. However, they have been catalogued in Table 22.3, along with references in which the details of their operation is explained.

22.6.2 Scanning Electron Microscope

The scanning electron microscope is probably the most widely used analysis instrument in the field of nanophotonics. A highly focused beam (typically 10–20 nm diameter) of 0.3–30 keV electrons is scanned across the surface of the sample. A 3D image of the sample surface can be formed by using backscattered electrons

Table 22.3 Analytical techniques used in nanotechnology

Technique	Acronym	Application	Reference
X-ray diffraction	XRD	Characterize crystal structure	[27] [28]
X-ray photoelectron spectroscopy	XPS (ESCA)	Surface composition and electronic states	[27]
Transmission electron microscopy	TEM	Structure of very thin specimens	[27, 28]
Scanning electron microscopy	SEM	Image of structure	[27, 28]
Low-energy electron diffraction	LEED	Surface structure	[28]
Reflection high-energy electron diffraction	RHEED	Monitor layer-by-layer MBE growth	[27]
Energy dispersive spectroscopy	EDS	Chemical composition of a microscopic area (used with SEM)	[27]
Electron microprobe	EMP	Surface layer composition	[28]
Scanning probe microscopy	SPM	3D real space images localized properties	[27]
Scanning tunneling microscopy	STM	Images of structure	[27]
Atomic force microscopy	AFM	Images of structure	[27]

and ejected secondary electrons. In addition surface properties can be characterized from measured cathodoluminescence and characteristic low-energy x-rays generated by the incident electrons. Resolution is limited by the electron beam diameter. The image of a DFB grating in Fig. 15.10b is a good example of a high-resolution SEM photomicrograph. It is important to note that SEM samples must be at least somewhat electrically conducting. Otherwise electronic charge will build up on the surface and deflect incident and secondary electrons. To avoid this problem, non-conducting samples can be coated with a thin layer of metal or graphite. Graphite sprays are commercially available for this purpose.

22.6.3 Reflection High-Energy Electron Diffraction

Another analysis tool that is widely used in the nanophotonics field is reflection high-energy electron diffraction (RHEED). MBE systems are usually fitted with a RHEED gun and screen so as to permit layer-by-layer analysis of the crystallinity of the grown layer. Thus, any problems can be detected and corrected at the earliest point in time. A high-energy (10–20 keV) beam of electrons is directed at a small angle (1–3°) toward the growth surface. Diffracted electrons are backscattered from the first few layers of the surface and the diffraction pattern is observed on a phosphorescent screen. From the diffraction pattern morphology and intensity one can determine layer thickness, crystalline symmetry and growth pattern. The measurements are made *in situ*, molecular layer by layer as the growth continues, so an

MBE film containing a particular number of molecular layers can be controllably produced by properly operating the shutters over the effusion sources. In this way multiple layers of different materials such as GaAs and GaInAsP can be grown with well-controlled thickness and composition.

22.7 Nanophotonic Devices

Nanophotonic technology has now been applied to most of the devices described in the previous chapters. The ultimate goals of this research are to produce devices with enhanced performance and to create PICs with smaller size and greater complexity. The work is still in its early stages, but many notable achievements have been reported.

22.7.1 Waveguides

A fundamental element of any PIC is the waveguide. Thus, it is not surprising that the fabrication of waveguides in photonic crystals is a topic that was given early investigation. It has been demonstrated that introduction of a line defect into a two-dimensional, slab photonic crystal can be used to create a waveguide. A line defect is created by adding or removing dielectric material along a line in the lattice of the PhC. This is sometimes called “defect doping” by analogy with impurity doping of a semiconductor. In “acceptor” doping the effective index at a lattice site is reduced by replacing some higher-index material with a lower-index material (e.g., increasing the size of air holes in a slab). In “donor” doping the effective index at a lattice site is increased by replacing some lower-index material with a higher-index material (e.g., reducing the size of air holes in a slab). A defect of this type, formed by eliminating a line of holes in a slab PhC, is shown in Fig. 22.15 [7]. In this case, the slab is a silicon-on-insulator (SOI) layer.

Light is confined to the waveguide in the x–y plane for states that exist at certain frequencies within the photonic band gap. Light is confined in the z-direction, vertical to the plane, by total internal reflection (TIR) due to the changes in refractive index at the surfaces of the slab. (Acceptor doping is not used to confine light in a line-defect waveguide of this type because it would be confined to the regions of lower index, i.e., the air, and confinement to the slab in the vertical direction would be lost.)

The dispersion diagram for this waveguide is shown in Fig. 22.16a. The light gray region marks the photonic band gap. The diagram reveals that this waveguide is multimode, with four propagating modes lying within the photonic band gap, and below the light line. The light line is established by the refractive indices of the cladding materials above and below the PhC slab. The modes below the light line ideally are not subject to propagation losses. Above the light line, in the dark gray region of the diagram, the spectrum becomes a continuum of states with resonances

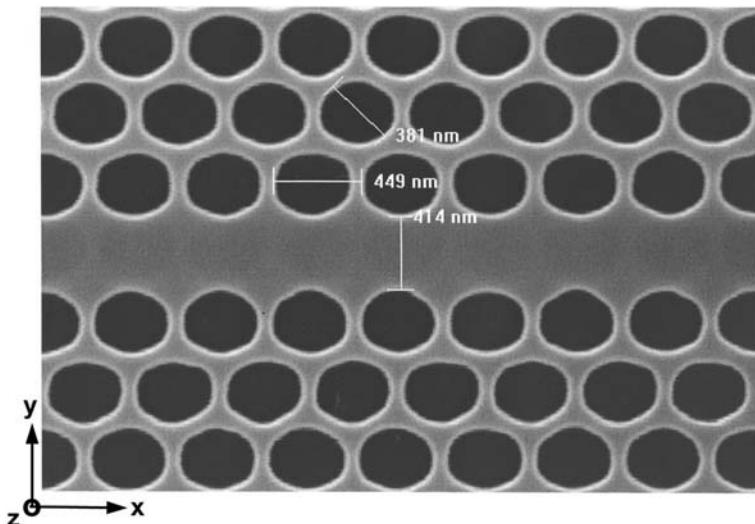


Fig. 22.15 An SEM photo of a line-defect waveguide in a slab PhC [7]

called quasi-guided modes. These modes correspond to leaky radiation modes of the waveguide that experience propagation losses. Fig. 22.16b shows the steady-state field distributions in the x–y plane of the four propagating modes calculated by the PWM. Two of the modes are even with respect to the horizontal mirror plane of the slab and two are odd.

In a conventional index-guiding waveguide, in which guiding is by TIR, single-mode propagation can be obtained by reducing the cross-sectional dimensions of the guide until all but one mode is cut off. However, that technique does not work well in PhC waveguides, in which the guiding of different modes is determined by interference effects and resonances rather than index variations. One must use a resonance effect to block the propagation of unwanted modes. For example, Baba et al. [29] have shown that adding a point defect to a line defect like that of Fig. 22.15, by removing a couple of additional holes on either side of the line, can produce a single-mode waveguide. They have reported that, if an additional defect is doped into the line defect, a mode with a resonant frequency at the cutoff range is strongly localized at the defect, while other modes overlapping with the waveguide band are radiated out through the waveguide. Thus a single propagating mode remains.

Line defect waveguides can be made to cross each other within the plane of the PhC slab without incurring excessive cross talk. For example, the structure shown in Fig. 22.17 has been shown by Johnson et al. [30] to exhibit nearly 100% throughput and 0% cross talk.

The calculated field pattern of a standard 90° crossing, shown in Fig. 22.17a, exhibits a significant amount of cross talk. The modified crossing of Fig. 22.17b,c is seen to be essentially free of cross talk.

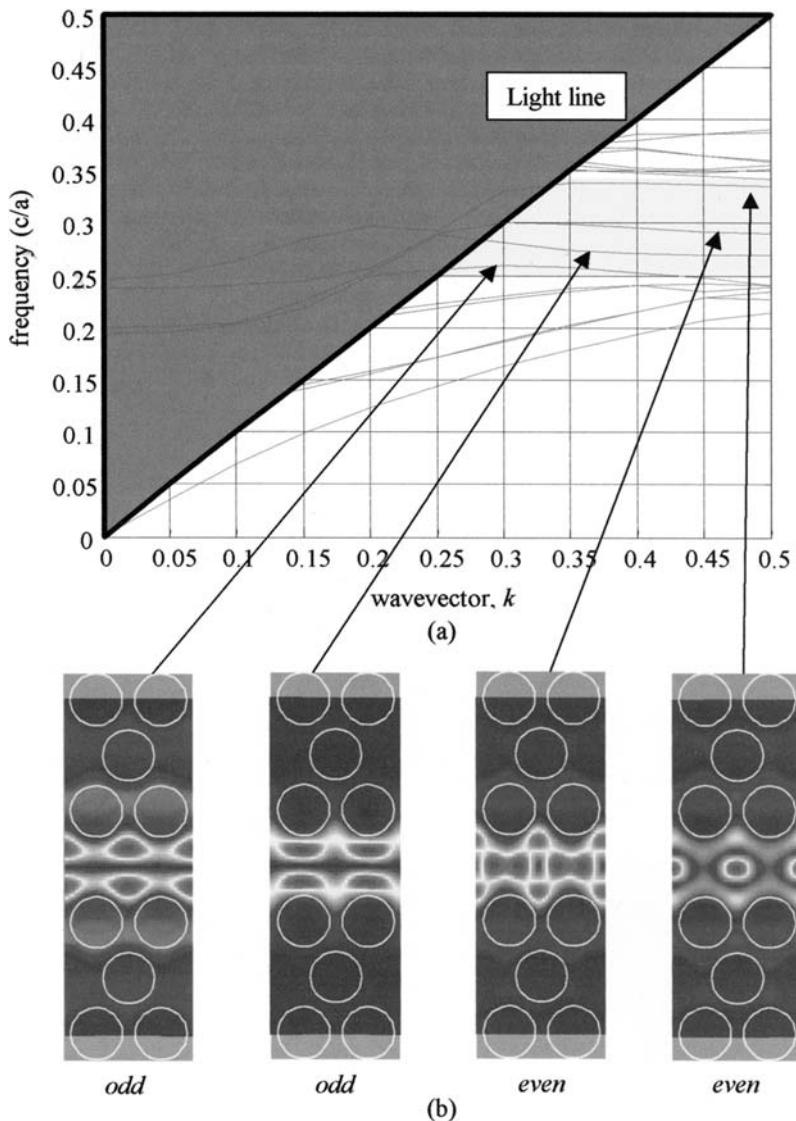


Fig. 22.16 Guided modes of a PhC single-line-defect waveguide: (a) dispersion diagram and (b) field profiles of the guided modes [7]

There are other types of optical waveguides in PhCs that are not based on line defects. Yariv et al. [31] have proposed a coupled-resonator optical waveguide (CROW) based on weak coupling between point-defect resonant cavities within a PhC. The dispersion relation of the CROW band can be described by a small coupling parameter κ , and the spatial characteristics of the CROW modes remain the

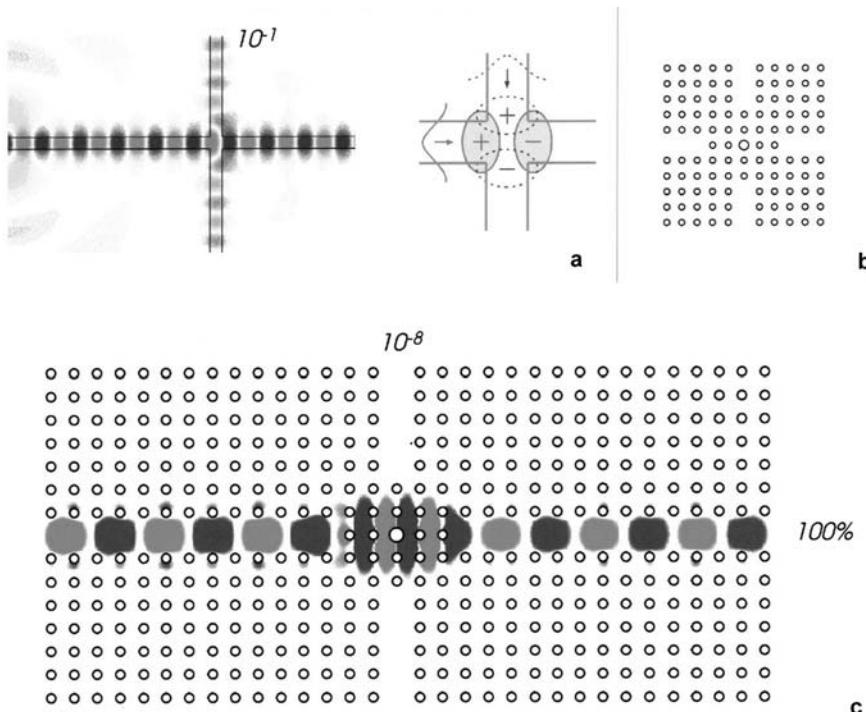


Fig. 22.17 Diagram of a line-defect-waveguide crossing: **a** perpendicular crossing; **b** modified crossing; and **c** field pattern in modified crossing [30]

same as those of the single-resonator high Q modes. CROWs can be utilized in constructing waveguides without cross talk. They also exhibit enhanced efficiency of second-harmonic generation [32]. It has been demonstrated that a section of CROW inserted into a line-defect waveguide alters its group velocity and positive/negative group velocity dispersion. This approach can be used to produce delay lines or dispersion compensators in PICs connected by linear-defect PhC waveguides [33]. The design of CROW delay lines, including calculations of delay, loss and bandwidth, have been described by Poon et al. [34]. Because CROWs can significantly decrease the group velocity of light they are sometimes called slow light structures. Crows may contain bends and branches, suggesting that efficient tunable CROW-based routers and switches may be realized [35]. In fact, a CROW Mach-Zehnder interferometer has already been demonstrated [36].

Another class of nanophotonics waveguides is nanowires. Nanowires can be defined as structures that have cross-sectional dimensions of tens of nanometers or less and any length. Because of the quantum mechanical effects that occur at these dimensions they are sometimes also called quantum wires. Many different types of nanowires exist, including metallic (e.g., Ni, Pt, Au), semiconducting (e.g., Si, InP, GaN, etc.) and insulating (e.g., SiO_2 , TiO_2).

Nanowire waveguides can be conveniently fabricated on silicon-on-insulator (SOI) substrates, which consist of a thin layer of silicon ($n=3.45$) on top of an oxide buffer layer ($n=1.45$). Either DUV, EUV or e-beam lithography is used to define waveguides with lateral core dimensions ~ 10 nm, and dry etching is used to produce waveguides surrounded by air on three sides and SiO_2 on the bottom. Thus $\Delta n > 2$ on all sides of the core and the required tight confinement is provided. Bogaerts et al. [37] have used this type of nanowire waveguide to make array waveguide gratings (AWGs), Mach-Zehnder lattice filters (MZLFs) and ring resonators. They used SOI wafers consisting of a 220-nm thin silicon layer on top of a 1- μm oxide cladding layer. They used DUV exposure at $\lambda = 248$ nm and directly etched into the top silicon layer with an ICP-RIE etch, using the resist as the etch mask. The buried oxide was not etched [38, 39]. Propagation losses in straight guides of this type have been measured as low as 0.24 dB/mm [38]. For a 90° bend with a radius of 3 μm , the excess loss per bend was determined to be 0.016 dB. Compact couplers [40] and array waveguides [41] have both been made based on silicon nanowire waveguides.

Metallic nanowires can also be used as waveguides at optical frequencies. In this case the optical frequency waves are actually carried by charged “particles” called surface plasmon-polaritons. A plasmon is a quasiparticle that results from the quantization of plasma oscillations, just as photons and phonons are quantizations of light and vibrational waves, respectively. In a plasma, ions are relatively slow moving, but the free-electron gas can oscillate at optical frequencies, producing plasmons. Plasmons at the surface of a metallic nanowire can interact strongly with photons of light to produce a polariton. Surface plasmon-polaritons can support optical frequency modes on a metallic stripe nanowire. The fundamental modes are similar to the symmetric and antisymmetric modes of a rectangular cross section slab waveguide. However, there are also higher-order modes, such as corner modes. Al-Bader [42] has calculated dispersion and loss characteristics for a silver stripe embedded in silicon. The method of calculation used was a full-vectorial finite-difference scheme for the coupled transverse components of the magnetic field. He found that mode effective indices for the fundamental modes were approximately 4.6 for a stripe thickness of 20 nm and width of 1 μm . A review of recent progress in the confining and guiding of electromagnetic energy at visible and near-infrared frequencies using surface plasmon excitations sustained by metallic nanostructures is given by Maier [43].

22.7.2 Couplers

A number of different types of couplers and splitters have been reported that efficiently couple lightwaves between linear-defect waveguides in PhCs. Fan et al. [44] have done numerical simulations of electromagnetic-wave propagation in waveguide branches in photonic crystals to identify structures with near-complete transmission, such as the one shown in Fig. 22.18.

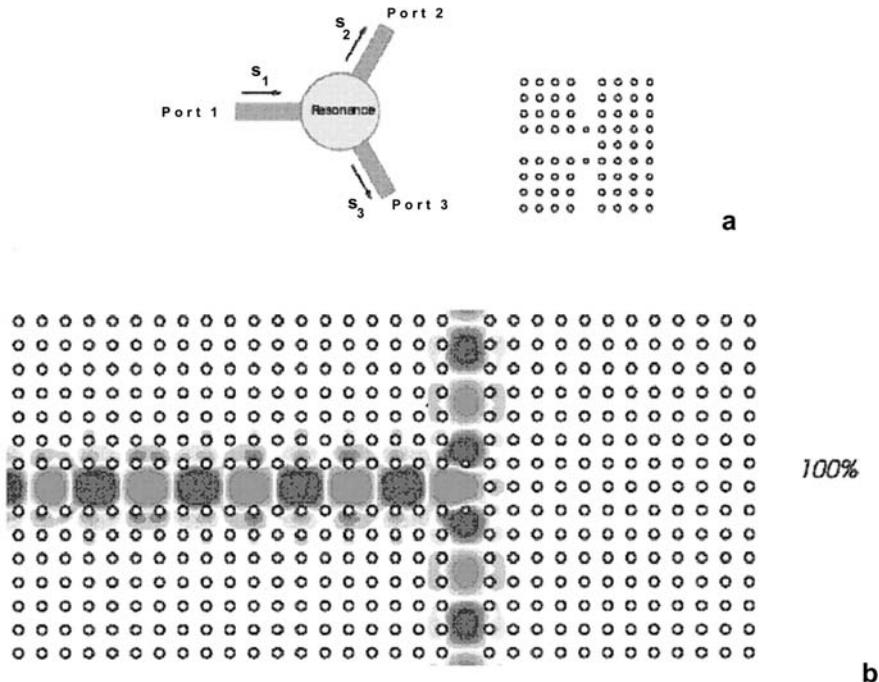


Fig. 22.18 Diagram of a line-defect-waveguide splitter: (a) 3-dB splitter structure and (b) field pattern I splitter [44]

The operation of this device depends on a tunneling process through localized resonant states between one-dimensional continuums [44]. This same mechanism has been employed in the design of an efficient channel-drop coupler [45].

Dual-channel directional couplers comprised of two closely spaced line-defect waveguides have been fabricated by Pustai et al. [46]. The devices were fabricated in the 260-nm-thick silicon device layer of a SOI wafer. A triangular PhC lattice of air holes was produced by e-beam lithography with a PMMA resist (200 nm thick) followed by RIE. The lattice constant, a , was 472 nm and the cylinder diameter, d , was 380 nm. The parallel-waveguide structure was made by masking a line of air holes from the etch to make a first waveguide, and masking a second line of air holes parallel to the first guide and separated from it by a single line of air holes. Couplers were made with three different between-guide spacings by changing the diameter of the air holes in the row separating the two coupled waveguides. This changed the amount of overlap of the evanescent tails and hence the coupling coefficient. The couplers were experimentally tested with a tunable laser with a range of 1260–1380 nm. For $d = 380$ nm, no coupling was observed over the 28- μm length of the coupler. (This demonstrates that these waveguides can be spaced at 360 nm without crosstalk.) For $d = 292$ nm, a majority of the light coupled for wavelengths in the range 1260–1281 nm, while for greater wavelengths the confinement was

better and the overlap of the mode tails was insufficient to produce significant coupling. This coupling effect was seen more strongly when the separation was further reduced by making $d = 232$ nm. Experimental results were in good agreement with theoretical simulations based on the PWM and FTDT methods.

The line-defect-waveguide couplers discussed so far have all coupled from one waveguide to another of similar dimensions. However, often it is necessary to couple a lightwave beam to a waveguide from the air, or from a waveguide of a different size. One way to accomplish this is to use a tapered coupler. The taper should be adiabatic in order to prevent mode conversion as lightwaves progress through it. An ideal structure would be one that adiabatically tapers in both the vertical and lateral directions and can be monolithically integrated with the PIC. Sure et al. [47] have described a method of producing 3D adiabatically tapered structures that involves writing a single grayscale mask in high-energy beam-sensitive (HEBS) glass with a high-energy electron beam. This is then used as a mask for the grayscale exposure of AZ4620 resist. The exposed resist is developed to produce continuous device profiles with a maximum and minimum height of 10 and 0.25 μm , respectively. The profiles are then transferred to a device layer of a SOI substrate by inductively coupled plasma etching. Experimentally determined coupling efficiency of 45% was demonstrated at 1.55 μm wavelength when coupling from air to a PIC waveguide 200 nm thick, and 75% was obtained when coupling to a 2- μm -thick guide.

Coupling to a nanoscale waveguide in a PhC has also been done by using modified versions of some of the couplers described in earlier chapters, such as prisms [48], gratings [49] and micro-lenses [50], as well as some new device geometries, such as the J Coupler [51].

22.7.3 Resonators

The waveguide splitter of Fig. 22.18 used a resonant structure to provide the necessary coupling from the main channel to the two secondary channels. And the addition of a single point defect to the waveguide of Fig. 22.15 was all that was necessary to form a resonant structure that blocked the higher-order modes and made the waveguide single mode. Such resonant structures are found frequently in PhC devices and PICs. By tradition, they are sometimes called “cavities” even though they lack conventional cavity form.

The similarities between a point-defect structure in a PhC and a traditional cavity resonator can be seen in Fig. 22.19. This PhC cavity was created by reducing the radius of the lattice site within the cavity by a factor of $\frac{1}{2}$ [7]. Like a conventional optical cavity it consists of a localized mechanical structure with recognizable “walls”. They both establish a time-invariant distribution of electromagnetic energy, or mode, and they both select a characteristic resonant frequency that is related to their dimensions and dielectric constant.

A PhC point-defect cavity can be tuned to a different frequency by changing either its dimensions or its dielectric constant. For example, Ripin et al. [52] have

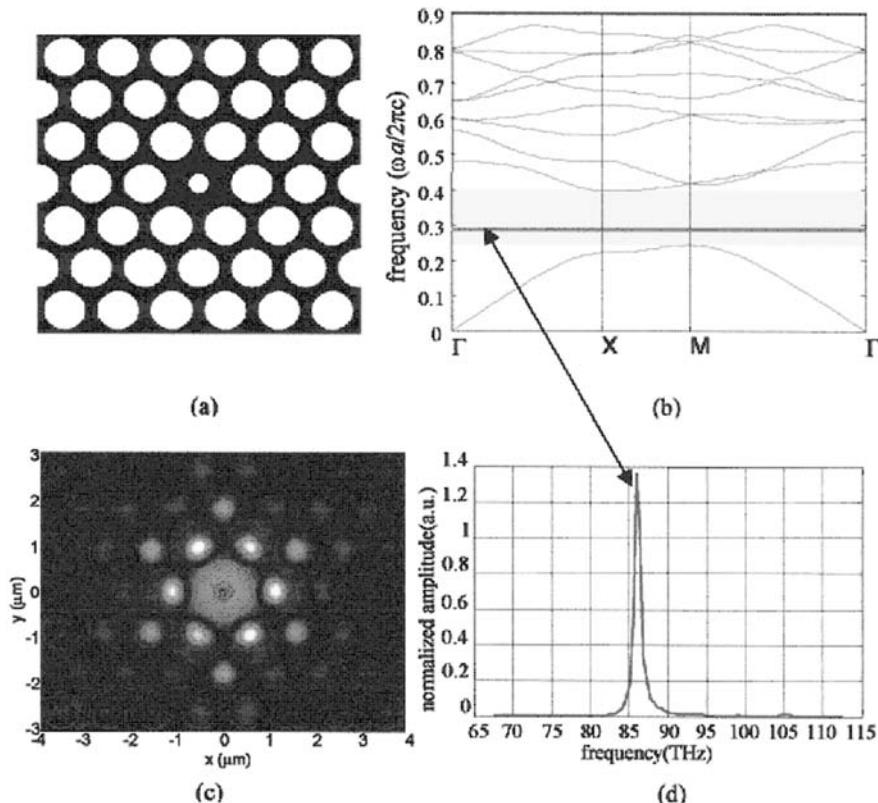


Fig. 22.19 (a, b, c, d) A photonic crystal point-defect cavity: (a) a donor-type point defect in a PhC lattice; (b) the dispersion where the state produced by the defect is marked by the *dark line*; (c) steady-state field profile of the cavity mode for TE polarization; and (d) the frequency response [7]

made the air-bridge resonator shown in Fig. 22.20. The PhC consists of a linear array of regularly spaced holes in a GaAs slab that is supported on Al_xO_y piers to form an air-bridge resonator. A microcavity is formed by varying the spacing, d , of two central holes. The resonant wavelength depends on the size of d . Quality factors as high as 360 were measured at wavelengths near $1.55 \mu\text{m}$, with modal volumes as small as $0.026 \mu\text{m}^3$.

Other types of PhC cavities with tunable resonant frequencies can also be made. For example, Nakagawa and Fainman [53] have made a tunable cavity based on modulation of the near-field coupling between two sub-wavelength periodic nanostructures embedded in a Fabry-Pérot cavity. The filter is composed of two planar dielectric mirrors forming a Fabry-Pérot cavity, with a field localization nanostructure affixed to the inner surface of each mirror. Each nanostructure consists of a transverse infinite sub-wavelength periodic grating with period $\Lambda = 0.6 \lambda_0$, where λ_0 is the intended resonant optical wavelength. A small air gap of $0.01\lambda_0$ between

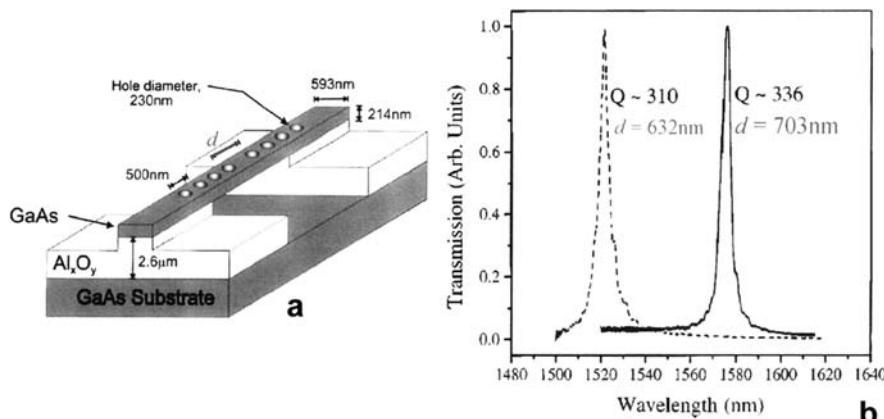


Fig. 22.20 (a, b) Air-bridge resonator: (a) diagram of structure and (b) transmission spectrum

the nanostructures allows near-field coupling between them. Either lateral or longitudinal mechanical displacement of the two nanostructures causes a change in the resonant frequency of the cavity.

22.7.4 Light Emitters

There is currently a great deal of interest in the topic of nanoscale light emitters. Most of this work has been done employing nanowires. GaN nanowires grown on silicon have been found to exhibit a high light output compared to that of bulk GaN [54, 55]. This improved performance is attributed to fewer defects, strains and impurities. Light-emitting diodes have been fabricated with these nanowires. Improved light emission efficiency has also been observed in GaN nanowires grown on sapphire [56]. Nanowire lasers also have been reported. Experimental near-field images of lasing nanowires provide a strong evidence that lasing occurs into the modes which are guided by the nanowire and propagate along its axis [57, 58]. Nanowire lasers with integrated electrodes have been intensity modulated [59]. Visible and ultraviolet nanoscale lasers consisting of single CdS or GaN nanowires were electric field using microfabricated electrodes. Studies of the electric field dependence in devices of different geometries indicated that modulation is due to an electroabsorption mechanism.

Not all nanoscale light emitters have been made in nanowires. Some have been made with quantum dots. Si quantum dot LEDs, imbedded in GaN thin layers, have exhibited high efficiency [60]. The external quantum efficiency of these devices was improved by inserting transparent doping layers between the Si quantum dot active layer and a transparent electrode layer, and by adding a multilayer microcavity DBR structure. A variety of quantum dot laser structures have been produced on both silicon and III-V material substrates. Mi et al. [61] have reported room temperature

operation of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ quantum dot lasers grown directly on Si substrates with a thin $0.5 \mu\text{m}$ GaAs buffer layer. Bakir et al. [62] have reported the fabrication and operation of a heterogeneous and compact “2.5 D” PhC microlaser with a single plane of InAs quantum dots as gain medium. The devices consisted of a top two-dimensional InP PhC slab, an SiO_2 bonding layer, and a bottom high-index-contrast Si/SiO_2 Bragg mirror deposited on a Si wafer. Choi et al. [63] have produced InAs, DBR, $1.55 \mu\text{m}$ wavelength, quantum dot lasers capable of being modulated at a data rate of 2.5 Gbps. Seven-stacked quantum dot layers with the dot density of $2 \times 10^{10}/\text{cm}^2$ were grown on an InP substrate with a grating structure and processed to be ridge waveguide-type laser diode. For further details of quantum dot light emitters, see, for example, Michler [64].

Nanoscale light emitters have also been made from various types of nanocrystals. A light-emitting diode structure based on Si nanocrystals formed by implantation into thermal oxide has been reported by Lalic and Linnros [65]. Iacona et al. [66] have made light-emitting diode structures based on Er-doped Si nanoclusters. Erbium doping results in the presence of an intense EL at $1.54 \mu\text{m}$ with a concomitant disappearance of a secondary 900-nm emission.

22.7.5 Photodetectors

For a complete nanoscale PIC, nano-photodetectors are required. In response to this requirement a number of different types have been developed. Zheng et al. [67] have made a nano-sized, p–n-junction, avalanche photodiode (NAPD). The device structure, shown in Fig. 22.21, consists of two portions, a photoelectron collector composed of an n-type floating cathode and an absorption region beneath it, combined with a nano-pillar which contains a charging layer, a multiplication region and a

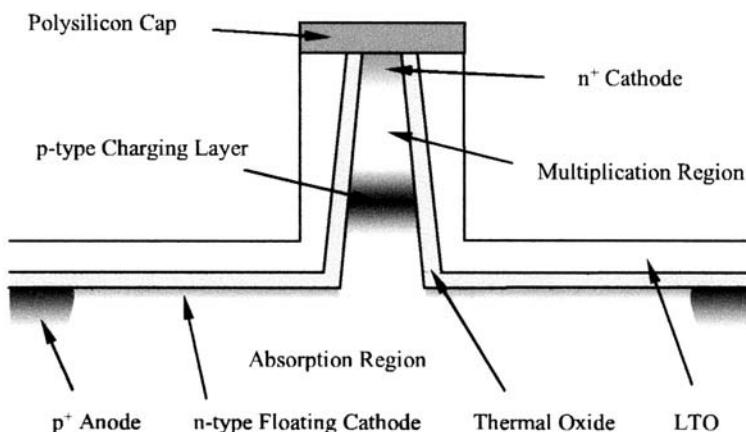


Fig. 22.21 Diagram of NAPD structure

heavily doped n-type cathode. The nano-pillar has a typical height of 800–1000 nm and a typical average diameter of 100–150 nm.

Simulations for the designed NAPD structure predicted a nearly constant avalanche gain over the reverse bias range of 15–30 V. Gain of 100 was predicted for a charging layer doping dose of $10^{14}/\text{cm}^2$. Experimental results were in good agreement with the predicted performance.

Nayfeh et al. [68] have produced thin-film, silicon, nanoparticle UV photodetectors by room temperature deposition of Si nanoparticle films on Si p-type substrates. The particle size was approximately 1 nm. The current–voltage characteristics indicated a photoconductor in series with a diode-like junction with a large increase in the forward current under UV illumination. In these devices, illumination by 365 nm wavelength light produced photocurrents $\sim 1 \text{ mA}$, while for visible light ($\lambda = 560 \text{ nm}$) photocurrent was $\sim 0.01 \text{ mA}$. A photodiode that is sensitive to UV wavelengths, but not visible wavelengths, is particularly useful in situations in which there is a high background level of visible light.

In Section 22.7.1 it was described how optical-frequency signals could be carried on a metallic nanowire waveguide, supported by charged surface plasmon–polaritons. A photodetector for this type of optical-wavelength signal preferably contains a nano-sized, metallic, resonant antenna. Such a photodiode has been developed [69–71]. Ishi et al. [69] have fabricated a Si Schottky photodiode that consists of an active area of 300 nm in diameter and a surface plasmon antenna. The diode exhibits an increase of the photocurrent by several decades compared to a similar device without a surface plasmon antenna. This result suggests an enhanced photogeneration of carriers via surface plasmon resonance. Since the opto-electronic signal conversion process occurs within a sub-wavelength scale, photodiodes with a surface plasmon antenna are inherently high-speed devices. The interconnection of devices within a PIC by metallic nanowire waveguides in conjunction with photodiodes of this type may provide an effective way to produce high-speed nanoscale PICs.

22.7.6 Sensors

A major field of application for nanophotonic devices is that of sensors. It is beyond the scope of this book to give a thorough review of all of the different types of nanophotonic sensors that have been developed, but a number of representative examples will be described. For additional information on this topic the reader is referred to the book by Andrews and Gaburro [72].

Nanophotonic sensors have been proposed for the detection of micro-damage in composite materials. Simulation models predict that a nanophotonic sensor attached to a composite bar experiences a significant change in its bandgap profile when damage is induced in the composite bar [73]. The sensor could be a PhC or a hollow-core photonic bandgap fiber. For a discussion of hollow-core photonic bandgap fibers, see, for example, Vincetti et al. [74] or Smith et al. [75].

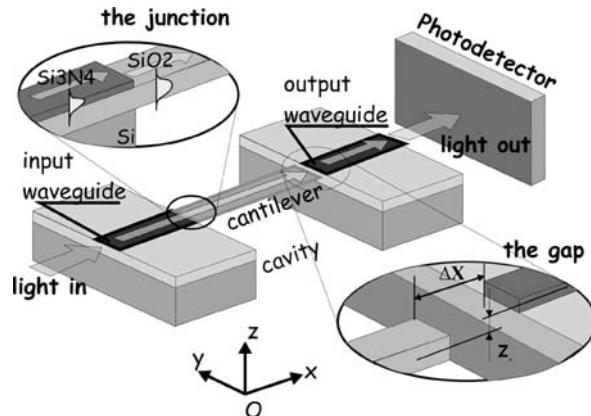
There is a class of nanophotonic sensors that is based on the interaction of the evanescent tail of a waveguided mode with the surrounding media and its effect on either the real or imaginary part of the waveguide mode effective index. Perturbations of the real part of the index change the mode's velocity, phase and wavelength. Perturbations of the imaginary part cause a change in losses and attenuation. All of these changes can be detected and measured to make a sensor. For example, an evanescent field sensor has been developed by Densmore et al. [76]. The device consists of a SOI photonic-wire waveguide incorporated into a Mach-Zehnder interferometer-based sensor. It can be used to monitor the properties of a homogeneous solution, making it useful as a biomedical transducer. Mach-Zehnder interferometer (MZI) sensors were fabricated on a SOI wafer with an Si-layer thickness of $0.26\text{ }\mu\text{m}$ and a buried oxide thickness of $2\text{ }\mu\text{m}$. A single-mode photonic-wire waveguide of rectangular cross section $0.26\text{ }\mu\text{m} \times 0.45\text{ }\mu\text{m}$ was patterned by e-beam lithography and RIE. In a solution of H_2O , the evanescent tail of the mode for $1.55\text{ }\mu\text{m}$ wavelength extended approximately $0.8\text{ }\mu\text{m}$ into the water. Changes in the effective index N_{eff} of the guided (TM) mode were measured as a function of changes in the index N_c of a sugar/water superstrate solution. The measured sensitivity was $\delta N_{\text{eff}} / \delta N_c = 0.31$, for $\lambda = 1.55\text{ }\mu\text{m}$.

Another type of evanescent tail sensor is the tapered-fiber waveguide. By heating a portion of a conventional step-index glass optical fiber and then putting it under controlled tension, a drawn-out reduced diameter section can be created. This section can act as a sensor because the thinned cladding and the reduced diameter core cause the evanescent tail of the mode to extend out beyond the cladding into the surrounding medium. Corres et al. [77] have made a nanophotonic version of this type of sensor by coating the reduced diameter section with a nanolayer of humidity-sensitive material. The tapered section typically had a waist diameter of $20\text{--}25\text{ }\mu\text{m}$ and a length of 1 mm. The original fiber was a standard single-mode communications fiber with $n_{\text{core}} = 1.4573$ and $n_{\text{clad}} = 1.450$. A humidity-sensitive nanofilm of poly(diallylmethylammonium chloride) (PDDA) and the polymeric Dye R-478 (Poly-R) was deposited by the electrostatic self-assembly (ESA) method [78]. A $20\text{ }\mu\text{m}$ waist fiber coated with a 275 nm polymer overlay was experimentally tested at a wavelength of 1310 nm . The change in the intensity transmitted through the fiber varied from 0 to 14 dB as the ambient humidity was increased from 75 to 100%.

Nanophotonic sensors can also be produced by using the techniques of micro-opto-electro-mechanical systems (MOEMS) described in Chapter 19. A sensor of this type has been produced by Zinoviev et al. [79], who have developed the micro-cantilever beam sensor shown in Fig. 22.22.

The principle of operation is based on the sensitivity of the coupling efficiency between two butt-coupled waveguides to their misalignment with respect to each other. The authors report that the cantilever deflection can be detected with a resolution of $18\text{ fm}/(\text{Hz})^{1/2}$, where the limiting factor is the shot noise of the photodetector. The device they fabricated featured an array of 20 independent waveguide channels, with $200\text{-}\mu\text{m}$ -long cantilevers, on a single chip of silicon. The cantilevers were 500 nm thick. Both the cantilevers and the waveguides were $40\text{ }\mu\text{m}$ wide.

Fig. 22.22 Schematic diagram of microcantilever sensor [79] © 2006 IEEE



The sensitivity of this type of device is comparable to one using the atomic force microscope (AFM) detection principle but can be used more conveniently in applications in which the parallel real-time monitoring of several responses on one chip is desired.

There is today a great need for biological sensors to be used in health sciences, environmental monitoring and biological warfare countermeasures. The nanoscale size of PhCs and their sensitivity to changes in the media surrounding them make them ideal choices for many of these types of applications. Biological substances may be sensed directly by measuring their dielectric properties, or indirectly by attaching a fluorescent compound to them and then exciting the fluorescent molecules. The former method is called “label-free detection” and the latter is called “enhanced fluorescence detection”. PhCs have been successfully employed in both of these detection methods [80–84].

Viruses typically have a size in the range of 20–200 nm [85]. This small size makes them difficult to detect in real time as required by biological warfare countermeasure applications. Two optical real-time detection schemes for nano-sized particles have been developed by Ignatovich et al. [86]. One of the methods involves measuring the optical gradient force exerted on a nanoparticle as it passes through a confined optical field, and the other method employs an interferometric scheme to detect the scattered field amplitude from a laser-irradiated particle.

For further information on nanoscale biological sensors the reader is referred to the books by Shi [87], Hoffmann [88] and Schultz et al. [89].

22.8 Future Projections for Integrated Optics and Nanophotonics

The technical feasibility and economic viability of optical fiber interconnectors have clearly been demonstrated in a wide variety of applications. The use of optical fiber

cable transmission systems is now widespread for interexchange links and under-sea cables, and has been extended to many businesses and homes to provide high-speed Internet access and data transmission. Digital data transmission rates on the order of tens of gigabits per second are commonplace, and rates up to several terabits per second will become common within the next few years in DWDM systems. Optical fiber cables have not yet replaced copper wire cables in the subscriber loop in many locations but the plans for that conversion are already being made. When it occurs, the public telephone network will be transformed into a vastly more flexible medium for the delivery of all sorts of new services including interactive high-speed data transmission, high-definition TV cable channels, video telephone, high-speed FAX, and security and medical alarm systems. Most households are already enjoying the benefits of data communication over the Internet and the World Wide Web.

With high-data-rate optical-fiber-cable transmission lines becoming well established, the need for optical integrated circuits has substantially grown. As a result, optical integrated circuits have ceased to be merely laboratory curiosities and have entered the marketplace. A number of companies are now marketing optical integrated circuits as off-the-shelf commercial products, and several companies are offering to produce user-designed PICs in a “foundry” type of service such as has been available for many years for electronic IC chips. Presently the commercially available PICs are relatively limited – laser diode chips, 2×2 optical switches, electro-optic phase modulators, Mach-Zehnder interferometer intensity modulators, multiplexers, demultiplexers, OADMs, couplers and splitters. However, these products have only been on the market for a relatively short time compared to electronic ICs. As systems engineers become more familiar with their use and their advantages, the demand for more sophisticated PICs will surely grow. As optical fibers are extended further into the subscriber loops of telecommunication systems the demand for large quantities of high-speed optical switches and couplers can best be met by producing them in PIC form.

The use of PICs and fibers as sensors is also a rapidly expanding field of application. So far most of the work has been confined to laboratory experiments and military applications, but a growing consumer market is evident. For example, the sensing and control functions in a modern automobile are becoming extremely complex and have been highly automated. Integrated optic sensors and signal processors with optical fiber interconnects can offer improved performance and reduced weight. Several major automobile manufacturers are already including fiber optic instrumentation systems and other photonic devices in their cars.

The combination of integrated optic technology with microwave devices and systems, particularly in regard to wireless communications as discussed in Sections 21.1 and 21.2, has the potential of providing another expanding market, with applications in data transmission, wireless Internet, cellphones, avionics, security systems and television distribution. The benefits of the Internet are already being provided to many wireless laptop and hand-held devices over combined optical fiber and microwave systems. Wireless photonic systems are also being widely used in security and surveillance applications.

The development of nanophotonics technology within the past few years has opened up an entirely new field of photonics in which the sizes of familiar devices have shrunk by orders of magnitude and new and unique devices, such as PhCs and microdot lasers, have been created. The potential applications for these devices are yet mostly unexplored. However, it is expected that their impact on our future lives will be significant and wide ranging, extending from engineering into physics, chemistry, biology and medicine. So many sectors of our society will be affected that many believe that we must investigate the cultural, ethical and educational impacts of nanotechnology with regard to potential risks as well as benefits. These aspects of nanotechnology have been considered in a number of books [90–93]. At this point in time, one cannot accurately predict the future of nanophotonics, but surely it will be exciting.

Problems

- 22.1 Describe how a photonic crystal guides light. What is the basic principle (mechanism) responsible for the guiding?
- 22.2 What are the advantages of a vertical cavity surface emitting laser (VCSEL) as compared to an end-face emitting laser?
- 22.3 An ideal lens of diameter = 5 cm and focal length = 1 m is used to focus light of wavelength = $1.55 \mu\text{m}$.
 - (a) What is the size of smallest object that the lens can resolve according to the Rayleigh criterion?
 - (b) If the same lens is used in a practical lithographic projection exposure system with resist factor $k = 0.80$, what is minimum feature size that can be produced? Hint: Assume $\text{NA} = D/2f$
- 22.4 What is the De Broglie wavelength of an electron traveling on a vacuum with a velocity $= 2 \times 10^7 \text{ m/s}$?
- 22.5 Describe two sputter-etching techniques that can be used to produce high-aspect-ratio structures in semiconductors and dielectric materials.
- 22.6 A plane wave is traveling in a two-dimensional rectangular photonic crystal lattice with lattice constant = 660 nm. What wavelength(s) of light will be blocked and not able to propagate?
- 22.7 What are the two growth methods used to produce layers in III-V semiconductors like GaAs and InP? Explain the special steps that must be taken in order to grow nanolayers as opposed to microlayers.

References

1. G.J. Parker, M.D.B. Charlton: Photonic crystals, Phys. World **13**, 29 (2000)
2. R. Dingle: *proc. 13th Int'l Conf, on Physics of Semiconductor*. F.G. Fumi (ed.) (North Holland. Amsterdam 1976)

3. S.J. Yoo, J.W. Lim, Y. Sung, Y.H. Jung, H.G. Choi: Fast switchable electrochromic properties of tungsten oxide nanowire bundles, *Appl. Phys. Lett.* **90**, 173126 (2007)
4. P. M. Petroff, A. Lorke, A. Imamoglu: Epitaxially self-assembled quantum dots, *Phys. Today* **54**, 94 (2001)
5. Y. Arai, T. Yano, S. Shibata: High refractive-index microspheres as optical cavity structure, *J. Sol-Gel Sci. Technol.* **32**, 189 (2004)
6. S.F. Tang, C.D. Chiang, P.K. Weng, Y.T. Gau, J.J. Luo, S.T. Yang, C.C. Shih, S.Y. Lin, S.C. Lee: High-temperature operation normal incident 256 × 256 InAsGaAs quantum-dot infrared photodetector focal plane array, *IEEE Photonics Technol. Lett.* **18**, 988 (2006)
7. D.M. Pustai: *Realizing functional two-dimensional Photonic crystal devices*, Ph. D. Dissertation (University of Delaware, Newark, DE 2004)
8. K.M. Leung, Y.F. Liu: Photon Band structures-the plane wave method, *Phys. Rev. B-Condens. Matter* **41**, 10188 (1990)
9. L. Liu, J.T. Liu: Photonic band structure in the nearly plane wave approximation, *Eur. Phys. J. B*, 381 (1999)
10. A. Taflove: *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, Inc., Boston, MA 1995)
11. D.R. Smith, S. Shultz, S.L. McCall, P.M. Platzmann: Defect studies in a 2-dimensional Periodic Photonic Lattice, *J. Mod. Opt.* **41**, 395 (1994)
12. H. Ibach, H. Lüth: *Solid-State Physics, An Introduction to Principles of Materials Science*, 2nd Ed. (Springer-Verlag, New York, Heidelberg, 1996)
13. C. Kittel, *Introduction to Solid State Physics* (Wiley: New York, 1996)
14. D.M. Pustai, A. Sharkawy, S.Y. Shi, D.W. Prather: Tunable photonic crystal microcavities, *Appl. Opt.* **41**, 5574 (2002)
15. M. Wilson, K. Kannagara, G. Smith, M. Simmons, B. Raguse *Nanotechnology: Basic Science and Emerging Technologies* (Chapman & Hall/CRC, Boca Raton, FL. 2002)
16. T. Fukui, R. Saito: International Symposium on GaAs and Related Compounds. Biarritz, France (1984).
17. I. Kim, D.G. Chang, P.D. Dapkus: *J. Cryst. Growth* **195**, 138 (1998)
18. S.A. Campbell: *The Science and Engineering of Microelectronic Fabrication* 2nd edn. (Oxford University Press, Oxford, New York, 2001) Chap. 9
19. J.E. Bjorkholm: EUV lithography-The successor to optical lithography?, *Intel Technol. J.* **Q3**, 1 (1998)
20. J.L. Vossen, W. Kern: *Thin Film Processes* (Academic Press, New York, 1978)
21. M. Koskenvuori, N. Chekurov, V-M, Airaksinen, I. Tittonen: Fast dry fabrication process with ultra-thin atomic layer deposited mask for released MEMS-devices with high electromechanical coupling, *Proc. Solid-State Sensors, Actuators and Microsystems Conference 2007* **10**, 501 (2007)
22. M.J. de Boer, J.G.E. Gardeniers, H.V. Jansen, E. Smulders, M-J. Gilde, G. Roelofs, J.N. Sasserath, M. Elwenspoek: Guidelines for etching silicon MEMS structures using fluorine high-density plasmas at cryogenic temperatures, *J. Microelectromech. Syst.* **11**, 385 (2002)
23. C. Chang, Y-F Wang, Y. Kanamori, J-J. Shih, Y. Kawai, C-K. Lee, K-C. Wu, M. Esashi: Etching submicrometer trenches by using the Bosch process and its application to the fabrication of antireflection structures, *J. Micromech. Microeng.* **15**, 580 (2005)
24. M. Puech, JM Thevenoud, N. Launay, N. Arnal, P. Godinat, B. Andrieu, JM. Gruffat: High productivity DRIE solutions for 3D-SiP and MEMS volume manufacturing, *J. Phys: Conf. Ser.* **34**, 481 (2006)
25. R. Bosch GmbH: US Patents 5501893, 6127273, 6214161, 6284148 (filed 1994-1999, issued 1996-2001)
26. S. Venkataraman, J. Murakowski, T.N. Adam, J. Kolodzey, D.W. Prather: Fabrication of high-fill-factor photonic crystal devices on silicon-on-insulator substrates, *J. Microlith., Microfab. Microsyst.* **2**, 248 (2003)

27. P.N. Prasad: *Nanophotonics* (Wiley-Interscience, New York, 2004) Chap. 7
28. S.M. Sze: *VLSI Technology*, 2nd edn. (McGraw Hill, New York, 1988) Chap. 12
29. T. Baba, D. Mori, K. Inoshita, Y. Kuroki: Light localizations in photonic crystal line defect waveguides, *IEEE J. Sel. Top. Quant. Electron.* **10**, 484 (2004)
30. S.G. Johnson, C. Manolatou, S. Fan, P.R. Villeneuve, J.D. Joannopoulos, H.A. Haus: Elimination of cross talk in waveguide intersections, *Opt. Lett.* **23**, 1855 (1998)
31. A. Yariv, Y. Xu, R.K. Lee, A. Scherer: Coupled-resonator optical waveguide: a proposal and analysis, *Opt. Lett.* **24**, 711 (1999)
32. Y. Xu, R.K. Lee, A. Yariv: Propagation and second-harmonic generation of electromagnetic waves in a coupled-resonator optical waveguide, *JOSA B*, **17**, 387 (2000)
33. W.J. Kim, W. Kuang, J. O'Brien: Dispersion characteristics of photonic crystal coupled resonator optical waveguides, *Opt. Express* **11**, 3431 (2003)
34. J.K.S. Poon, J. Scheuer, Y. Xu, A. Yariv: Designing coupled-resonator optical waveguide delay lines, *JOSA B* **21**, 1665 (2004)
35. S.V. Boriskina: Spectral engineering of bends and branches in microdisk coupled-resonator optical waveguides, *Opt. Express* **15**, 17371 (2007)
36. A. Martinez, A. Griol, P. Sanchis, J. Marti: Mach Zehnder interferometer employing coupled-resonator optical waveguides, *Opt. Letters* **28** (6) 405 (2003)
37. W. Bogaerts, P. Dumon, D. VanThourhout, D. Taillaert, P. Jaenen, J. Wouters, S. Beckx, V. Wiaux, R.G. Baets: Compact wavelength-selective functions in silicon-on-insulator photonic wires, *IEEE J. Sel. Top. Quant. Electron.* **12**, 1394 (2006)
38. W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. VanThourhout, Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology, *J. Lightwave Technol.* **23**, AOI (2005)
39. W. Bogaerts, V. Wiaux, D. Taillaert, S. Beckx, B. Luyssaert, P. Bienstman, R. Baets: Fabrication of photonic crystals in silicon-on-insulator using 248-nm deep UV lithography, *IEEE J. Sel. Top. Quant. Electron.* **8**, 926 (2002).
40. D. Dai, S. He: Optimization of ultracompact polarization-insensitive multimode interference couplers based on Si nanowire waveguides, *IEEE Photonics Technol. Lett.* **18**, 2017 (2006)
41. D. Dai, S. He: Ultrasmall overlapped arrayed-waveguide grating based on Si nanowire waveguides for dense wavelength division demultiplexing, *IEEE J. Sel. Top. Quant. Electron.* **12**, 1301 (2006)
42. S.J. Al-Bader: Optical transmission on metallic wires-fundamental modes. *IEEE J. Quant. Electron.* **40**, 325 (2004)
43. S.A. Maier: Plasmonics: metal nanostructures for subwavelength photonic devices, *IEEE J. Sel. Top. Quant. Electron.* **12**, 1214 (2006)
44. S. Fan, S.G. Johnson, J.D. Joannopoulos, C. Manolatou, H.A. Haus, Waveguide branches in photonic crystals, *J. Opt. Soc. Am. B* **18**, 162 (2001)
45. S. Fan, P.R. Villeneuve, J.D. Joannopoulos, H.A. Haus: Channel drop tunneling through localized states, *Phys. Rev. Lett.* **80**, 960 (1998)
46. D.M. Pustai, A. Sharkawy, S. Shi, G. Jin, J. Murakowski, D.W. Prather: Characterization and analysis of photonic crystal photonic waveguides, *J. Microlith., Microfab. Microsyst.* **2**, 292 (2003)
47. A. Sure, T. Dillon, J. Murakowski, C. Lin, D. Pustai, D.W. Prather: Fabrication and characterization of three-dimensional silicon tapers, *Opt. Lett.* **27**, 1601 (2002)
48. N. Ryusuke, M. Takashi, T. Yasuhiro, J. Piprek: Novel nano-defect measurement method of SOI wafer using evanescent light, *SPIE Proc.* **6013**, 60130 N.1 (2005)
49. F. Van Laere, T. Claes, J. Schrauwen, S. Scheerlinck, W. Bogaerts, D. Taillaert, L. O'Faolain, D. Van Thourhout, R. Baets: Compact focusing grating couplers for silicon-on-insulator integrated circuits, *Photonics Technol. Lett.* **19**, 1919 (2007)
50. H.M. Presby, C.A. Edwards: Near 100% efficient fiber microlenses, *Electron. Lett.* **28**, 582 (1992).

51. D.W. Prather, J. Murakowski, S. Shi, S. Venkataraman, A. Sharkawy, C. Chen, D. Pustai: High-efficiency coupling structure for a single-line-defect photonic-crystal waveguide, Opt. Lett. **27**, 1601 (2002)
52. D.J. Ripin, K.Y. Lim, G.S. Petrich, P.R. Villeneuve, S. Fan, E.R. Thoen, J.D. Joannopoulos, E.P. Ippen, L.A. Kolodziejski: Photonic band gap airbridge microcavity resonances in GaAs/AlxOy waveguides, J. Appl. Phys. **87**, 1578 (2000) Fig. 22.22, Copyright 2000. Reprinted with permission from American Institute of Physics.
53. W. Nakagawa, Y. Fainman: Tunable optical nanocavity based on modulation of near-field coupling between subwavelength periodic nanostructures, IEEE J. Sel. Top. Quant. Electron. **10**, 478 (2004)
54. J.B. Schlager, N.A. Sanford, K.A. Bertness, J.M. Barker, A. Roshko and P.T. Blanchard: Polarization-resolved photoluminescence study of individual GaN nanowires grown by catalyst-free MBE, Appl. Phys. Lett. **88**, 213106 (2006)
55. K.A. Bertness, N.A. Sanford, J.M. Barker, J.B. Schlager, A. Roshko, A.V. Davydov, I. Levin: Catalyst-free growth of GaN nanowires. J. Electron. Mater. **35**, 576 (2006)
56. D-H. Kim, C-O. Cho, J. Kim, H. Jeon, T. Sakong, J. Cho, Y. Park: Extraction efficiency enhancement in GaN-based light emitters grown on a holographically nano-patterned sapphire substrate, Proceedings of the 2005 IEEE Quantum Electronics and Laser Science Conference, **2**, 1268 (2005)
57. J.C. Johnson, H.J. Choi, K.P. Knutsen, R.D. Schaller, P. Yang, R.J. Saykally: Single gallium nitride nanowire lasers, Nature Mater. **1**, (2002).
58. X. Duan, Y. Huang, R. Agarval, C.M. Lieber: Single-nanowire electrically driven lasers, Nature **421**, 241 (2003).
59. A.B. Greytak, C.J. Barrelet, Y. Li, C.M. Lieber: Semiconductor nanowire laser and nanowire waveguide electro-optic modulators, Appl. Phys. Lett. **87**, 151103 (2005)
60. K.S. Cho, N.-M. Park, T.-Y. Kim, K.H. Kim, G.Y. Sung, J.H. Shin: High efficiency visible electroluminescence from silicon nanocrystals embedded in silicon nitride using a transparent doping layer, Appl. Phys. Lett. **86**, 071909 (2005)
61. Z. Mi, P. Bhattacharya, J. Yang, K.P. Pipe: Room-temperature self-organised In_{0.5}Ga_{0.5}As quantum dot laser on silicon, Electron. Lett. **41**, 742 (2005)
62. B.B. Bakir, C. Seassal, X. Letartre, P. Regreny, M. Gendry, P. Viktorovitch, M. Zussy, L. Di Cioccio, J.-M. Fedeli, Room-temperature InAs/InP quantum dots laser operation based on heterogeneous “2.5 D” photonic crystal, Opt. Express **14**, 9269 (2006)
63. B.S. Choi, E.D. Sim, C.W. Lee, J.S. Kim, H.-S. Kwack, D.K. Oh: 1.55 um quantum dot laser for 2.5 Gbps operation, Proc. NSTI Bio-Nanotechnol. Conf., Santa Clara, CA (2007)
64. P. Michler (ed.): *Single Quantum Dots Fundamentals, Applications and New Concepts*, Top. Appl. Phys. **90** (Springer, New York, 2003)
65. N. Lalic, J. Linnros: Light emitting diode structure based on Si nanocrystals formed by implantation into thermal oxide. J. Lumin **80**, 263 (1999)
66. F. Iacona, A. Irrera, G. Franzo, D. Pacifici, I. Crupi, M. Miritello, C.D. Presti, F. Priolo: Silicon-based light-emitting devices: properties and applications of crystalline, amorphous and Er-doped nanoclusters, IEEE J. Sel. Top. Quant. Electron. **12**, 1596 (2006)
67. X. Zheng, A.L. Lane, B. Pain, T.J. Cunningham: Modeling and fabrication of a nano-multiplication-region avalanche photodiode, Proc. NASA Science Technology Conference, NSTC, College Park, MD, 2007
68. O.M. Nayfeh, S. Rao, A. Smith, J. Therrien, M.H. Nayfeh: Thin-film silicon nanoparticle UV photodetector, IEEE J. Sel. Top. Quant. Electron. **16**, 1924 (2004)
69. T. Ishi, J. Fujikata, K. Makita, T. Baba, K. Ohashi: Si nano-photodiode with a surface plasmon antenna, Jpn. J. Appl. Phys. **44**, L364 (2005)
70. J. Fujikata, T. Ishi, D. Okamoto, K. Nishi, K. Ohashi: Highly efficient surface-plasmon antenna and its application to Si nano-photodiode, Proc. 19th Annual Meeting of the IEEE Lasers and Electro-Optics Society, Montreal, Quebec (2006)

71. K. Nishi, J. Fujikata, T. Ishi, D. Okamoto, K. Ohashi: Development of nanophotodiodes with a surface plasmon antenna, Proc. 20th Annual Meeting of the IEEE Lasers and Electro-Optics Society, Lake Buena Vista, FL (2007)
72. D.L. Andrews, Z. Gaburro, (eds.), *Frontiers in Surface Nanophotonics*, Springer Series in Optical Sciences **133** (Springer, New York, 2007)
73. I. El-Kady, M.M.R. Taha: Nano Photonic sensors for microdamage detection: an exploratory simulation, Proc. IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, Hawaii (2005)
74. L. Vincetti, M. Maini, F. Poli, A. Cucinotta, S. Selleri: Numerical analysis of hollow core photonic band gap fibers with modified honeycomb lattice, *Opt. Quant. Electron.* **38**, 903 (2006)
75. C.M. Smith, N. Venkataraman, M.T. Gallagher, D. Müller, J.A. West, N.F. Borrelli, D.C.A. Koch, K.W. Koch: Low-loss hollow-core silica/air photonic
76. A. Densmore, D.-X. Xu, P. Waldron, S. Janz, P. Cheben, J. Lapointe, A. Delâge, B. Lamontagne, J.H. Schmid, E. Post: A silicon-on-insulator photonic wire based evanescent field sensor. *IEEE Photonics Technol. Lett.* **18**, 2520 (2006)
77. J.M. Corres, J. Bravo, I.R. Matias, F.J. Arregui: Nonadiabatic tapered single-mode fiber coated with humidity sensitive nanofilms. *IEEE Photonics Technol. Lett.* **18**, 935 (2006)
78. A. Ulman: Formation and structure of self-assembled monolayers. *Chem. Rev.* **96**, 1533 (1996)
79. K. Zinoviev, C. Dominguez, J.A. Plaza, V.J.C. Bustos, L.M. Lechuga: A novel optical waveguide microcantilever sensor for the detection of nanomechanical forces, *J. Lightwave Technol.* **24**, 2132 (2006)
80. L.L. Chan, S. Gosangari, K. Watkin, B.T. Cunningham: A label-free photonic crystal biosensor imaging method for detection of cancer cell cytotoxicity and proliferation, *Apoptosis* **12**, 1061 (2007)
81. B. Lin, P.Y. Li, B.T. Cunningham: A label-free biosensor-based cell attachment assay for characterization of cell surface molecules, *Sens. Actuators B* **114** (2006)
82. P.C. Mathias, N. Ganesh, L.L. Chan, B.T. Cunningham: Combined enhanced fluorescence and label-free biomolecular detection with a photonic crystal surface, *Appl. Opt.* **26**, 2351 (2007)
83. N. Ganesh, W. Zhang, P.C. Mathias, E. Chow, J.A.N.T. Sooares, V. Malyarchuk, A.D. Smith, B.T. Cunningham: Enhanced fluorescence emission from quantum dots on a photonic crystal surface, *Nature Nanotechnol.* **2**, 515 (2007)
84. N. Ganesh, I.D. Block, B.T. Cunningham: Near ultraviolet-wavelength photonic-crystal biosensor with enhanced surface-to-bulk sensitivity ratio, *Appl. Phys. Lett.* **89**, 023901 (2006)
85. C.A. Tidona, G. Darai, C. Buuchen-Osmond: *The Springer Index of Viruses* (Springer-Verlag Berlin, 2002)
86. F.V. Ignatovich, D. Topham, L. Novotny: Optical detection of single nanoparticles and viruses, *IEEE J. Sel. Top. Quant. Electron.* **12**, 1292 (2006)
87. D. Shi (ed.): *NanoScience in Biomedicine* (Springer-Verlag, New York, 2008)
88. K.-H. Hoffmann (ed.): Functional Micro-and Nanosystems Proceedings of the 4th Caesarium, Bonn, June 16-18, 2003 (Springer-Verlag, New York, 2004)
89. J. Schultz, M. Mrksich, S.N. Bhatia, D.J. Brady, A.J. Ricco, D.R. Walt, C.L. Wilkins, (eds.), *Biosensing-International Research and Development*, (Springer-Verlag, New York, 2006)
90. H. Brune, H. Ernst, A. Grunwald, W. Grunwald (eds.): *Nanotechnology: Assessment and Perspectives* (Springer-Verlag, Berlin, Heidelberg, 2006)
91. N. Cameron, E. Mitchell: *Nanoscale: Issues and Perspectives for the Nano Century* (Wiley, New York, 2007)
92. R.E. Hester, R. Harrison (eds.): *Nanotechnology: Consequences for Human Health & the Environment* (Springer-Verlag, New York, 2007)
93. P.M. Boucher: *Nanotechnology: Legal Aspects* (CRC Press, Boca Raton, Florida, 2008)

Index

A

Absorption, 54, 64, 110–116, 178–181, 221, 233, 249, 332, 336, 345, 363–366
free-carrier, 112–116, 279, 349–350
interband, 110, 225–227

Acoustic beam deflector, response time, 210

Acoustic frequency modulator, 331

Acoustic-optic effect, 210–214

Acoustic-optic figure of merit, 210

Acoustic transducer, interdigitated, 210, 211–213, 215
multi-periodic, 212, 213
multiple array, 212
phased array, 213
single-periodic, 210, 211

Acoustic wave, 210

Advantages of integrated optic systems, 2–8

Analog-to-digital converter, 428–429

Analyzer, 177

Angle, Bragg, 188, 189, 209, 212

Angle, coupling, 23, 122, 123, 136, 140

Annealing, 60, 64–66, 360

Array waveguides, 93–95, 444

Attenuation
coefficient (α), 108, 119, 125–126
in OICs, 119, 124, 125, 263
in optical fibers, 4, 285, 287

Avalanche photodiode, 351–353
reliability, 353

B

Backward wave, 306–308

Balanced bridge modulator, 428

Band-edge light emission, 232

Bandwidth
acoustic modulator, 213–214, 410
modulator, 177
optical fiber, 6
relation to switching time, 172

Beam deflector, 208–214

Blazed grating, 140

Boltzmann statistics, 232

Boundary conditions, 19, 31–36, 42, 280, 282

Bragg angle, 188, 203, 205, 212

Bragg condition, 306

Bragg effect, 188, 189

Bragg reflection coupling efficiency, 303–305
grating, effect of lattice damage, 310
grating, fabrication, 309–315
grating, location, 310–313
higher-order gratings, 306–307
mode frequency spacing, 309
wavelength dependence, 303–305

Bulk acousto-optic modulator, 205

C

Cable television (CATV) systems, 440

Cadmium sulfide, 294

Cadmium sulfide selenide, 294

Cantilever beams, 406, 411–413

Carrier confinement, 283–284

Carrier diffusion time, 366

Carrier drift time, 366

Carrier lifetime, 267

Carrier recombination, 226, 227

Carrier trapping, 366

Catastrophic failure of laser diodes, 295

Cavities, 491

Charge-coupled device (CCD), 425

Chemically assisted ion beam etching (CAIBE), 309

Cleaved mirrors, 235

Coherent radiation, 236

Commercially available OICs, 10

Conductive oxides, 351

Confinement factor(d/D), 266, 283

- Corrugated grating, 303, 308
 Coupled-mode theory, 157–160, 305
 Coupled-resonator optical waveguide (CROW), 489–490
Coupler, 491–493
 3dB, 155
 air beam to optical fiber, 130
 beam to waveguide, 130–131, 153–154
 butt-coupled ridge waveguides, 164
 confluent, 164
 direct focus (“end-fire”), 130
 directional, 183
 dual-channel, 154, 155–157, 160–163, 362
 dual-channel, fabrication, 162, 163,
 166, 167
 end-butt, 131–135
 end-fire, 122
 fiber to waveguide, 145–149
 grating, 139, 141–142
 prism, 25, 122, 123, 136–139
 tapered, 143
 transverse, 130–135
 waveguide to waveguide, 153–169
Coupling
 coefficient, 137, 138, 162, 197, 305
 efficiency, 123, 130, 131, 143, 146, 305
 length, 185
 loss, 123, 146
 synchronous (coherent), 153–154, 155,
 157–160, 181
Critical angle, 190
Cross talk, 4, 191
CROW, *see* Coupled-resonator optical waveguide (CROW)
Crystal momentum, 221
Current “kinks”, 290
Current trends in integrated optics, 423–447
- D**
Damage threshold level, 295
Deep reactive-ion etch (DRIE) processes, 484
Defect doping, 487
Delay time, turn-on, 342
Depletion layer photodiode, 345–350, 354
Detector
 dark current, 346
 high-frequency cutoff, 367
 linearity, 367
 proton-bombarded, 360–362
 quantum efficiency, 347, 351, 365
 spectral response, 355, 392
Detector array, 418
Differential quantum efficiency, 296
- Diffract theory**, 208–209
Diffused dopants, 57–58, 284
Diffusion length, electron, 279
Direct bandgap, 227, 228
Directional coupler tree, 8
Direct modulation of laser diodes, 325–337,
 339–340
Dispersion-compensated fiber, 445
Dispersive optical element (DOE), 417–418
Divergence angle, 279
Doppler velocimeter, 429–430
Dry etching, 482
- E**
Efficiency of light emission, 237
Einstein coefficients, 233
E versus k diagram, 224, 225
Electro-absorption, 363–366
Electromagnetic interference (EMI), 3
Electrons and photons
 confinement of, 471–472
 properties of, 469–471
Electron transitions
 direct and indirect, 224, 225
 interband and intraband, 225
Electro-optical effect, 174
Electro-optic tensor, 174
Emitter/detector terminal, 364–365
Enhanced fluorescence detection, 499
Epitaxial growth, 65–72, 278–279, 288,
 356–359
Erbium-doped fiber, 272
Erbium doped fiber amplifier (EDFA),
 440–444
Erbium-doped materials, 9
Evanescence mode tail, 166, 294–297
External quantum efficiency definition, 228
Extinction coefficients, 41, 109, 281–284
- F**
Fabry-Perot etalon, 133, 235
Fiber Bragg gratings (FBG), 98–100
Fiber lasers, 271–273
Fiber-optic communication system, 2, 418
Fiber-optic sensors, 388
Field effect transistor (FET), 337, 436
Finite-difference time-domain (FDTD)
 method, 474–475
Franz-Keldysh effect, 178, 179, 363–364
- G**
Gain equalization, 271
Gain guiding, 290
Gallium arsenide, 9, 10, 23, 478

Gallium phosphide, 65, 228
Gaussian beam, 140
Geodesic lens, 424
Germanium, 356
Gold masking, 160
Goos–Hänchen shift, 29, 30
Graded-index fiber, 6, 286
Gradual degradation of laser diodes, 296
Grating
 fabrication, 141–142, 310
 periodicity, 140, 141
Guard ring structure, 353

H

Heterodyning, 425
High-voltage applications, 4
High-voltage sensors, 408, 409
History of integrated optics, 1, 2
Hole burning, 290
Holographic optical element (HOE), 417
Holographic photoresist exposure, 142
Hybrid integration, 9, 11, 13, 134,
 387–388

I

Image converter, 21
Impurity banding, 231
Index ellipsoid, 174
Index of refraction, effective, 29, 67, 79
Indirect bandgap, 225
“Infinite melt” epitaxy, 69
Injection pumping, 250
Insertion loss, 173, 387, 391
Integrated lenses, 417
Integrated optical detectors, 345–368
Integrated optics and nanophotonics, future
 projections for, 499–500
Interaction length, 153, 155, 156, 185, 201
Interdigitated transducer pattern, 211–212
Internal quantum efficiency definition, 229
Internet, 1
Inverted population, definition, 233
Ion
 exchange, 59
 implantation, 60–62, 77, 78
 implantation system, 61
 migration, 59
Ion beam etching, 141, 162, 309, 313
Ion-bombardment-enhanced etching, 163
Isolation, 173

K

Kerr effect, 174

L

Label-free detection, 491
Laser
 amplifier, 265, 267
 confined-field, 247, 250, 277–298
 for fiber-optic applications, 285–287
 monolithically integrated, 291–292
 pumping, 228–230, 234, 253–254
 quantum well, 377–384
 in quaternary materials, 287
 semiconductor, 241–258
 separate confinement heterostructure
 (SCH), 311, 313
 single heterostructure (SH), 278–279
 single-mode, 288–291, 317, 417
 stripe geometry, 288
 threshold, 235–237, 243–247, 252
 tunnel-injection, 250–252
 threshold conditions, 250–252
Laser diode, 241–249, 291–298
 control of emitted wavelength, 285–288
 direct modulation, 325–337,
 339–340
 frequency response, 329
 double heterostructure, 287
 effective gain, 247
 emission spectrum, 292, 381, 382
 extrapolated lifetime, 296
 gain coefficient, 245
 heterojunction, 250, 278–285
 long wavelength, 287–288
 loss coefficient, 245
 microwave modulation, 336–337
 mode spectrum, 243
 output power and efficiency, 248–249
 package, 336, 337
 power flow diagram, 245
 range finder, 255
 rate equations, 326
 reliability, 295
 spatial divergence, 244
 threshold conditions, 252
 transverse spatial energy distribution, 245
Laser, distributed Bragg reflection (DBR),
 313–315
 effective cavity length, 314
 longitudinal mode spacing, 314
Laser, distributed feedback (DFB),
 2, 303–324
 emission linewidth, 317
 laterally coupled, 312, 313
 oscillation condition, 309
 in quaternary materials, 287

- Laser, distributed feedback (*cont.*)
 threshold current density, 317–319
 wavelength selectability, 315
 wavelength stability, 316–319
- Lattice constants, 11
- Lattice mismatch, 11, 70–72
- LEAD devices, 365, 366
- Light emitters, 495–496
- Light-emitting diodes, 229
- Linewidth of light emission, 246, 272, 316, 317
- Link gain, 454
- Lithium niobate, 9, 53, 59, 85, 87, 93, 97, 194, 195
- Lithium tantalate, 174, 189
- Local oscillators, 425
- Loss, 107–128
- M**
- Mach-Zehnder interferometer (MZI), 498
- Magnetron sputtering, 483
- Maximum deviation, 172
- Maximum modulation depth (extinction ratio), 326
- Maxwell's wave equation, 18, 25, 30, 33, 282
- MBE, *see* Molecular beam epitaxy (MBE)
- Mesa photodiode, 349
- Metalorganic chemical vapor deposition (MOCVD), 72
- Metalorganic vapor phase epitaxy (MOVPE), 478–479
- Micro-opto-electro-mechanical devices (MOEMs or MEMs), 403–422
- Microscopic model for light generation and absorption, 221–226
- Microwave carrier generation, 459–463
- Microwave modulators, 337
- Microwave/optical integrated circuits, 337, 387, 423–436
- Microwave over fiber, 458
- M lines, 24, 123
- Mode instability, 290
- Model dispersion, 6, 271
- Modes, 17–31
 air radiation, 19, 27
 definition of, 17
 in dual-channel coupler, 155
 experimental observation of, 21–25
 guided, 19, 26
 in laser diodes, 237, 241–249, 264
 in planar waveguides, 18, 19, 33–39
 in rectangular waveguides, 39–48
 substrate radiation, 19, 27
- Mode spacing, 248
- Modulation
 amplitude, 325–328
 depth, 171
 direct control of laser diode current, 325–337, 339–340
 factors limiting maximum frequency, 332–337
 frequency, 330–331
 index, Bragg, 188
 index, Raman-Nath, 207
 at microwave frequencies, 331–337
 at mm-wave frequencies, 452
 pulse, 328–330
 pulse position, 329
 pulse width, 329
- Modulator
 acousto-optic, 201–220
 acousto-optic, Bragg Type, 204–208, 424–427
 acousto-optic, intermode coupling, 207
 acousto-optic, Raman-Nath type, 203
 bulk, 173, 193
 COBRA, 185
 dual-channel electro-optic, 181–200
 electro-absorption, 178–181
 electro-optic, 171–200, 187
 electro-optic, Bragg type, 188, 189
 electro-optic, Mach-Zehnder type, 193, 194
 electro-optic, total internal reflection, 190
 external power, 191
 intensity, 178
 phase, 175
 polarization, 177
 power consumption, 173
 waveguide, 175–181
- Molecular beam epitaxy (MBE), 71–72, 392, 478
- Monolithically integrated detectors, 337–339
- Monolithic integration, 9–10, 134, 291–295, 337
- MOVPE, *see* Metalorganic vapor phase epitaxy (MOVPE)
- Multiplexing, 3, 6, 414
- Multiplication factor, 352
- N**
- Nanomachining, 481–485
- Nanophotonic devices
 couplers, 491–493
 light emitters, 495–496

- photodetectors, 496–497
resonators, 493–495
sensors, 497–499
- Nanophotonics
dimensions, 469
electrons and photons, properties of, 469–471
integrated optics and nanophotonics, future projections for, 499–500
nanophotonic devices, *see* Nanophotonic devices
nanostructures, characterization and evaluation of
available tools, 485
scanning electron microscope, 485–486
nanostructures, fabrication of, 477
metalorganic vapor phase epitaxy, 478–479
molecular beam epitaxy, 478
nanomachining, 481–485
nanoscale lithography, 479–481
- photonic crystals
classes of, 472–473
comparison of electrons in semiconductor crystals to photons in, 473–477
- photons and electrons, confinement of, 471–472
- Nanophotonic sensors, 498
- Nanoscale lithography, 479–481
- Nanostructures, benefits of, 470
- Nanostructures, characterization and evaluation of
available tools, 485
reflection high-energy electron diffraction, 186–487
scanning electron microscope, 485–486
- Nanostructures, fabrication of, 477
metalorganic vapor phase epitaxy, 478–479
molecular beam epitaxy, 478
nanomachining, 481–485
nanoscale lithography, 479–481
- Noncentrosymmetric crystal, 174
- Normalization, 157
- Number of resolvable spots, 208
- O**
- OIC sensors, 425, 426, 432–433
- Optical add drop multiplexer (OADM), 413, 414
- Optical bistability, 339
- Optical disk readhead, 430–432
- Optical feedback, 233, 291, 303
- Optical integrated circuit (OIC), 1, 2, 8–15, 423–436, 452, 464
- Optical pumping, 228, 294
- Optical switch, 175, 191
- Optical tunneling, 137, 153–160, 294
- Optoelectronic integrated circuit (OEIC), 394–398, 436–437
- Opto-microwave monolithic integrated circuit (OMMIC), 452
- Organosilicon films, 57
- Overlap integral, 130
- P**
- Parasitic reactance, 332–336
- Phase constant difference ($\Delta\beta$), 159, 160
- Phased array, 213
- Phased array antenna driver, 438–439
- Phase matching, 136, 139, 162, 202
- Phonon, definition, 228
- Photocurrent, 346, 350
- Photodetectors, 496–497
- Photodiode
metal-semiconductor-metal, 354–355
signal-to-noise ratio, 367
spectral response, 355–366
- Photoelastic effect, 201
- Photoelastic tensor, 201
- Photomultiplication factor, 352
- Photon absorption, conservation of energy and momentum, 228, 229
- Photon, definition, 224
- Photon emission, conservation of energy and momentum, 228, 229
- Photonic crystals
classes of, 472–473
comparison of electrons in semiconductor crystals to photons in, 473–477
- Photonic integrated circuits (PIC), 469
- Photon
lifetime, 327
mass, 225
- Photoresist masking, 75, 142, 162
- PIC, *see* Photonic integrated circuits (PIC)
- Piezoelectric micrometer, 133
- Piezoelectric transducer, 206, 210–211
- Plane wave method (PWM), 474
- Plasma frequency, 62
- Pockels effect, 174
- Polarization sensitivity, 186
- Polyimide, 89
- Polymer waveguides, 12, 53, 57, 85–89, 164
- Polystyrene, 87–89

- Praseodymium-doped fiber amplifiers, 265
P
 Prism
 materials, 141
 spacing, 136
 Propagation constant, 18, 28, 117, 153
 Proton bombardment, 63, 64, 160, 288,
 360–363
- Q**
 Quantum confined Stark effect (QCSE),
 387–388
 Quantum photoconductivity, 392
 Quantum well, 377
 avalanche photodiodes, 392
 detectors, 392–393
 laser, 380–384
 modulators and switches, 384–391
 Quantum wires, 490
 Quartz substrate, 206
 Quasi-Fermi levels, 234
 Quasi-guided modes, 488
 Quaternary materials, 10, 65, 70, 180, 265,
 287, 346, 357
- R**
 Radiation loss, 116–120
 Raman optical amplifiers, 263–264
 Rayleigh criterion, 108, 480
 Ray optics, 25, 26–28
 Relaxation oscillation, 329
 Repeater spacing, 440, 441
 Resonant cavity, 236, 392
 RF spectrum analyzer, 423–426
 Ridge waveguides, 164
 Reactive-ion etching (RIE), 484
 Reflection high-energy electron diffraction
 (RHEED), 186–487
 Resonators, 493–495
 RHEED, *see* Reflection high-energy electron
 diffraction (RHEED)
 RIE, *see* Reactive-ion etching (RIE)
- S**
 Scanning electron microscope,
 485–486
 Scattering loss, 107–110, 123–125
 Schottky-barrier contact, 351
 Schottky-barrier photodiode, 350, 351
 Self-aligned couplers, 147
 Self electro-optic effect devices (SEED),
 393–394
 Self sustained oscillations, 329
 Sellmeier equation, 67
- Semiconductor optical amplifiers,
 265–269
 Sensors, 497–499
 Silicon nitride, 55
 Single-mode optical fibers, 286
 Sidebar boat, 69
 Slow light structures, 490
 Spatial harmonics, 140
 Spiking resonance, 329
 Spontaneous emission of light, 226–232
 Sputtering, 53–55
 ion-beam, 55
 plasma discharge, 54
 reactive, 55
 Standing wave, 242
 Stepped $\Delta\beta$ reversal, 185, 186
 Stimulated emission or light, 232, 233
 Stop cleaving, 292
 Strain, 375
 Strain-optic tensor, 201
 Stress, 404
 Substitutional dopant atoms, 57–62
 Substrate materials for OICs, 8
 Super-radiance radiation, 236
 Surface acoustic wave (SAW), 210–211
 Surface plasmon–polaritons, 491
 Switching speed, 172, 191
 Switch matrix, 10
- T**
 Tantalum pentoxide, 80, 132, 206
 Telecommunications systems, 1, 3,
 439–447
 Temperature sensor, 432–433
 Ternary materials, 10, 65, 287
 Thermal effects in lasers, 295–298
 Thermal equilibrium, 227, 234
 Thermalization, 227
 Thin-film deposition, 54–58
 from solutions, 50
 Thin membranes, 405, 408–411
 Thulium-doped fiber amplifiers, 271
 Torsion plates, 407, 414–415
 Total internal reflection, 25, 137,
 190–191
 Transient multimode excitation, 332
 Transition rate equation, 233
 Transverse junction stripe (TJS) laser,
 292, 293
 Traveling wave electrodes, 265, 266
 Trends in optical telecommunications,
 439–444
 Tunneling current density, 251

V

Vacuum-vapor deposition, 53
Variance, 109
Velocity approach, 117–118
Vertical-cavity-surfaceemitting laser (VCSEL), 339

W

Waveguide
asymmetric, 38–39
branching, 164, 165, 295
carrier-concentration-reduction, 62–65, 162
channel, 40–45
curved, 117, 125, 126
cutoff conditions, 29, 37–39, 63, 80
depletion layer, 279, 280
electro-optic, 73–74

fabrication methods, 53–84, 161–162
loss measurement, 122–123
photodiode, 345, 349
planar, 18, 19, 33–39, 53–79
rectangular, 39–48, 94
strip-loaded, 45–48
symmetric, 36–38

Wavelength incompatibility, 65
Wireless network, 451, 452

X

X-ray lithography, 481

Z

Zinc oxide, 207
Zinc selenide, 58, 65
Zinc telluride, 65