ELSEVIER

SECOND EDITION

Rajiv Ramaswami

Kumar N. Sivarajan

Optical Networks

A PRACTICAL PERSPECTIVE

RESPECT COPYRIGHT
ELIMINATE PIRACY

This edition is for sale in Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka and designated countries in South-East Asia through Elsevier (Singapore) Pte. Ltd. Sale and purchase of this book outside of these countries is unauthorised by the publisher.

Nokia of America Corporation
IPR2023-00681, Ex. 1016
Ex. 1016.1

Optical Networks

A Practical Perspective

Second Edition

Ex. 1016.2

The Morgan Kaufmann Series in Networking
Series Editor, David Clark, MIT

Optical Networks: A Practical Perspective, 2e
Rajiv Ramaswami and Kumar N. Sivarajan

Telecommunications Law in the Internet Age
Sharon K. Black

Internet QoS: Architectures and Mechanisms
Zheng Wang

TCP/IP Sockets in Java: Practical Guide for Programmers
Michael J. Donahoo and Kenneth L. Calvert

TCP/IP Sockets in C: Practical Guide for Programmers
Kenneth L. Calvert and Michael J. Donahoo

Multicast Communication: Protocols, Programming, and Applications
Ralph Wittmann and Martina Zitterbart

MPLS: Technology and Applications
Bruce Davie and Yakov Rekhter

High-Performance Communication Networks, 2e
Jean Walrand and Pravin Varaiya

Computer Networks: A Systems Approach, 2e
Larry L. Peterson and Bruce S. Davie

Internetworking Multimedia
Jon Crowcroft, Mark Handley, and Ian Wakeman

Understanding Networked Applications: A First Course
David G. Messerschmitt

Integrated Management of Networked Systems: Concepts, Architectures, and their Operational Application
Heinz-Gerd Hegering, Sebastian Abeck, and Bernhard Neumair

Virtual Private Networks: Making the Right Connection
Dennis Fowler

Networked Applications: A Guide to the New Computing Infrastructure
David G. Messerschmitt

Modern Cable Television Technology: Video, Voice, and Data Communications
Walter Ciciora, James Farmer, and David Large

Switching in IP Networks: IP Switching, Tag Switching, and Related Technologies
Bruce S. Davie, Paul Doolan, and Yakov Rekhter

Wide Area Network Design: Concepts and Tools for Optimization
Robert S. Cahn

Practical Computer Network Analysis and Design
James D. McCabe

Frame Relay Applications: Business and Technology Case Studies
James P. Cavanagh

For further information on these books and for a list of forthcoming titles, please visit our Web site at *www.mkp.com.*

Optical Networks

A Practical Perspective

Second Edition

Rajiv Ramaswami
Kumar N. Sivarajan

Morgan Kaufmann Publishers
An imprint of Elsevier

Optical Networks, 2/e
Ramaswami & Sivarajan

Morgan Kaufmann Publishers
An Imprint of Elsevier
340 Pine Street, Sixth Floor,
San Francisco, CA 94104-3205, USA

© 2002 by Academic Press

Original ISBN : 1-55860-655-6

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means—electronic or mechanical, including photocopy, recording, or any information storage and retrieval system—without permission in writing from the publisher.

First Printed in India 2004
Reprinted 2006

Indian Reprint ISBN-13 : 978-81-8147-207-6
Indian Reprint ISBN-10 : 81-8147-207-1

This edition is for sale in Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka and designated countries in South-East Asia through Elsevier (Singapore) Pte. Ltd. Sale and purchase of this book outside of these countries is unauthorised by the publisher.

Published by Elsevier, a division of Reed Elsevier India Private Limited,
Shri Pratap Udyog, 274, Captain Gaur Marg, Sriniwaspuri,
New Delhi-110065 (India).

Printed and bound at
Saurabh Printers Pvt. Ltd., A-16, Sector IV, NOIDA 201 301.

To our parents.

Foreword

by Paul E. Green, Jr.
Director, Optical Network Technology (retired)
Tellabs, Inc.

If anyone needed evidence of the rapid pace of innovation in WDM optical telecommunications, a year-by-year perusal of the innovations reported in the journals and in the activities of start-up companies in the optical space would provide all needed persuasion. While the field is old—some fifteen years to be exact—it is also young in the sense that almost nothing in the technology has reached steady-state maturity. New optical bands are being opened up, unprecedented bitrates are being carried per wavelength over all-optical distances that were not thought possible not long ago, new forms of fiber are being introduced, and point-point links are now being organized by their owners into rings and other real network structures, using new optical switching and routing technologies.

Thus, it is a most welcome development to see that Ramaswami and Sivarajan have reorganized and updated their outstanding 1998 traversal of this fascinating and important field and have done so in a way that accurately portrays what is going on. The first edition reflected the authors' backgrounds, straddling the academic and industrial research community on one hand and the school of commercial hard knocks on the other, and it is still true that this new book provides a view of optical communication that is lacking in other frequently updated volumes in this field, whose authors lack the authority possessed by Ramaswami and Sivarajan.

At this juncture, there is a pause in the headlong rush to provide ever more bandwidth to everybody, and the reasons are not just economic. The bottleneck between the gigabit environment on everyone's desktop and the terabit environment of the telco and cable fiber infrastructure is a consequence of the failure of the available "last mile" communication technologies to seriously meet the challenge of providing

Ex. 1016.7

enough bandwidth. Between now and the next edition of this book, one may hope that fiber last mile solutions will finally begin to proliferate, building on the technical ideas captured here.

There can be no doubt that fiber optic communication can only play an ever widening, if invisible, role in our lives, and once again, I cannot recommend the book highly enough to anyone interested in the technical underpinnings.

Ex. 1016.8

Foreword to the First Edition

by Paul E. Green, Jr.
Director, Optical Network Technology (retired)
Tellabs, Inc.

Not too many years ago, whenever one wanted to send messages effectively, there were really only two choices—send them by wire or send them by radio. This situation lasted for decades until the mid-1960s, when the fiber optics revolution began, quietly at first, and then with increasing force as people began to appreciate that sending pulses of light through tiny strands of glass wasn't so crazy after all. This revolution is now in full cry, with 4000 strand miles of fiber being installed per day, just in the United States alone. Fiber has been displacing wire in many applications, and gradually it is emerging as one of the two dominant Cinderella transmission technologies of today, wireless being the other. One of these (wireless) goes anywhere but doesn't do much when it gets there, whereas the other (fiber) will never go everywhere but does a great deal indeed wherever it reaches. From the earliest days of fiber communication, people realized that this simple glass medium has incredible amounts of untapped bandwidth capacity waiting to be mined, should the day come when we would actually need it, and should we be able to figure out how to tap it. That day has now come. The demand is here and so are the solutions.

This book describes a revolution within a revolution, the opening up of the capacity of the now-familiar optical fiber to carry more messages, handle a wider variety of transmission types, and provide improved reliabilities and ease of use. In many places where fiber has been installed simply as a better form of copper, even the gigabit capacities that result have not proved adequate to keep up with the demand. The inborn human voracity for more and more bandwidth, plus the growing realization that there are other flexibilities to be had by imaginative use of the fiber, have led people to explore all-optical networks, the subject of this book.

Ex. 1016.9

Such networks are those in which either wavelength division or time division is used in new ways to form entire network structures where the messages travel in purely optical form all the way from one user location to another.

When I attempted the same kind of book in 1993, nobody was quite sure whether optical networking would be a roaring success or disappear into the annals of "whatever happened to . . ." stories of technology that had once sounded great on paper, but that had somehow never panned out in the real world. My book (*Fiber Optic Networks,* Prentice Hall) spent most of its pages talking about technology building blocks and lamenting their limitations since there was little to say about real networks, the architectural considerations underlying them, and what good they had ever done anybody.

In the last four years, optical networking has indeed really happened, essentially all of it based on wavelength division multiplexing, and with this book Ramaswami and Sivarajan, two of the principal architects of this success, have redressed the insufficiencies of earlier books such as mine. Today, hundreds of millions of dollars of wavelength division networking systems are being sold annually, major new businesses have been created that produce nothing but optical networks, and bandwidth bottlenecks are being relieved and proliferating protocol zoos tamed by this remarkably transparent new way of doing networking; what's more, there is a rich architectural understanding of where to go next. Network experts, fresh from the novelties of such excitements as the Web, now have still another wonderful toy shop to play in. The whole optical networking idea is endlessly fascinating in itself—based on a medium with thousands of gigabits of capacity yet so small as to be almost invisible, transmitters no larger than a grain of salt, amplifiers that amplify vast chunks of bandwidth purely as light, transmission designs that bypass 50 years of hard-won but complex coding, modulation and equalization insights, network architectures that subsume many functions usually done more clumsily in the lower layers of classical layered architectures—these are all fresh and interesting topics that await the reader of this book.

To understand this new networking revolution within a revolution, it is necessary to be led with a sure hand through territory that to many will be unfamiliar. The present authors, with their rare mixture of physics and network architecture expertise, are eminently qualified to serve as guides. After spending some time with this book, you will be more thoroughly conversant with all the important issues that today affect how optical networks are made, what their limitations and potentialities are, and how they fit in with more classical forms of communication networks based on electronic time division. Whether you are a computer network expert wondering how to use fiber to break the bandwidth bottlenecks that are limiting your system capabilities, a planner or implementer trying to future-proof your telephone network, a teacher planning a truly up-to-date communication engineering curriculum, a student

Ex. 1016.10

looking for a fun lucrative career, or a midcareer person in need of a retread, this volume will provide the help you need.

The authors have captured what is going on and what is going to be going on in this field in a completely up-to-date treatment unavailable elsewhere. I learned a lot from reading it and expect that you will too.

Contents

Foreword ix

Foreword to the First Edition xi

Preface xxvii

1 Introduction to Optical Networks 1

1.1 Telecommunications Network Architecture 3
1.2 Services, Circuit Switching, and Packet Switching 6
1.2.1 The Changing Services Landscape 9
1.3 Optical Networks 10
1.3.1 Multiplexing Techniques 12
1.3.2 Second-Generation Optical Networks 14
1.4 The Optical Layer 16
1.5 Transparency and All-Optical Networks 24
1.6 Optical Packet Switching 26
1.7 Transmission Basics 28
1.7.1 Wavelengths, Frequencies, and Channel Spacing 28
1.7.2 Wavelength Standards 30
1.7.3 Optical Power and Loss 31
1.8 Network Evolution 32
1.8.1 Early Days—Multimode Fiber 33
1.8.2 Single-Mode Fiber 35

Ex. 1016.12

1.8.3 Optical Amplifiers and WDM 37
1.8.4 Beyond Transmission Links to Networks 39
Summary 40
Further Reading 41
References 42

I Technology 47

2 Propagation of Signals in Optical Fiber 49
2.1 Light Propagation in Optical Fiber 50
2.1.1 Geometrical Optics Approach 50
2.1.2 Wave Theory Approach 55
2.2 Loss and Bandwidth 65
2.2.1 Bending Loss 67
2.3 Chromatic Dispersion 68
2.3.1 Chirped Gaussian Pulses 69
2.3.2 Controlling the Dispersion Profile 74
2.4 Nonlinear Effects 76
2.4.1 Effective Length and Area 77
2.4.2 Stimulated Brillouin Scattering 79
2.4.3 Stimulated Raman Scattering 80
2.4.4 Propagation in a Nonlinear Medium 81
2.4.5 Self-Phase Modulation 83
2.4.6 SPM-Induced Chirp for Gaussian Pulses 87
2.4.7 Cross-Phase Modulation 89
2.4.8 Four-Wave Mixing 90
2.4.9 New Optical Fiber Types 93
2.5 Solitons 98
2.5.1 Dispersion-Managed Solitons 100
Summary 101
Further Reading 102
Problems 103
References 104

3 Components 107
3.1 Couplers 108
3.1.1 Principle of Operation 110
3.1.2 Conservation of Energy 111
3.2 Isolators and Circulators 112
3.2.1 Principle of Operation 113

3.3 Multiplexers and Filters 115
3.3.1 Gratings 118
3.3.2 Diffraction Pattern 122
3.3.3 Bragg Gratings 123
3.3.4 Fiber Gratings 126
3.3.5 Fabry-Perot Filters 130
3.3.6 Multilayer Dielectric Thin-Film Filters 133
3.3.7 Mach-Zehnder Interferometers 135
3.3.8 Arrayed Waveguide Grating 139
3.3.9 Acousto-Optic Tunable Filter 143
3.3.10 High Channel Count Multiplexer Architectures 148
3.4 Optical Amplifiers 151
3.4.1 Stimulated Emission 152
3.4.2 Spontaneous Emission 153
3.4.3 Erbium-Doped Fiber Amplifiers 154
3.4.4 Raman Amplifiers 159
3.4.5 Semiconductor Optical Amplifiers 161
3.4.6 Crosstalk in SOAs 165
3.5 Transmitters 165
3.5.1 Lasers 166
3.5.2 Light-Emitting Diodes 176
3.5.3 Tunable Lasers 178
3.5.4 Direct and External Modulation 186
3.5.5 Pump Sources for Raman Amplifiers 190
3.6 Detectors 192
3.6.1 Photodetectors 192
3.6.2 Front-End Amplifiers 197
3.7 Switches 199
3.7.1 Large Optical Switches 201
3.7.2 Optical Switch Technologies 207
3.7.3 Large Electronic Switches 215
3.8 Wavelength Converters 216
3.8.1 Optoelectronic Approach 217
3.8.2 Optical Gating 218
3.8.3 Interferometric Techniques 220
3.8.4 Wave Mixing 223
Summary 224
Further Reading 225
Problems 226
References 232

4 Modulation and Demodulation 239
4.1 Modulation . . . 239
4.1.1 Signal Formats . . . 240
4.2 Subcarrier Modulation and Multiplexing . . . 242
4.2.1 Clipping and Intermodulation Products . . . 243
4.2.2 Applications of SCM . . . 245
4.3 Spectral Efficiency . . . 245
4.3.1 Optical Duobinary Modulation . . . 246
4.3.2 Optical Single Sideband Modulation . . . 248
4.3.3 Multilevel Modulation . . . 249
4.3.4 Capacity Limits of Optical Fiber . . . 249
4.4 Demodulation . . . 250
4.4.1 An Ideal Receiver . . . 252
4.4.2 A Practical Direct Detection Receiver . . . 253
4.4.3 Front-End Amplifier Noise . . . 254
4.4.4 APD Noise . . . 255
4.4.5 Optical Preamplifiers . . . 255
4.4.6 Bit Error Rates . . . 258
4.4.7 Coherent Detection . . . 263
4.4.8 Timing Recovery . . . 265
4.4.9 Equalization . . . 266
4.5 Error Detection and Correction . . . 267
4.5.1 Reed-Solomon Codes . . . 270
4.5.2 Interleaving . . . 271
Summary . . . 272
Further Reading . . . 273
Problems . . . 274
References . . . 279

5 Transmission System Engineering 283
5.1 System Model . . . 283
5.2 Power Penalty . . . 284
5.3 Transmitter . . . 287
5.4 Receiver . . . 288
5.5 Optical Amplifiers . . . 289
5.5.1 Gain Saturation in EDFAs . . . 290
5.5.2 Gain Equalization in EDFAs . . . 292
5.5.3 Amplifier Cascades . . . 293
5.5.4 Amplifier Spacing Penalty . . . 294

5.5.5 Power Transients and Automatic Gain Control 296
5.5.6 Lasing Loops 298
5.6 Crosstalk 299
5.6.1 Intrachannel Crosstalk 299
5.6.2 Interchannel Crosstalk 301
5.6.3 Crosstalk in Networks 303
5.6.4 Bidirectional Systems 303
5.6.5 Crosstalk Reduction 305
5.6.6 Cascaded Filters 307
5.7 Dispersion 308
5.7.1 Chromatic Dispersion Limits: NRZ Modulation 309
5.7.2 Chromatic Dispersion Limits: RZ Modulation 311
5.7.3 Dispersion Compensation 314
5.7.4 Polarization-Mode Dispersion (PMD) 320
5.8 Fiber Nonlinearities 323
5.8.1 Effective Length in Amplified Systems 323
5.8.2 Stimulated Brillouin Scattering 325
5.8.3 Stimulated Raman Scattering 326
5.8.4 Four-Wave Mixing 329
5.8.5 Self-/Cross-Phase Modulation 333
5.8.6 Role of Chromatic Dispersion Management 335
5.9 Wavelength Stabilization 335
5.10 Design of Soliton Systems 336
5.11 Design of Dispersion-Managed Soliton Systems 338
5.12 Overall Design Considerations 341
5.12.1 Fiber Type 341
5.12.2 Transmit Power and Amplifier Spacing 342
5.12.3 Chromatic Dispersion Compensation 343
5.12.4 Modulation 343
5.12.5 Nonlinearities 344
5.12.6 Interchannel Spacing and Number of Wavelengths 344
5.12.7 All-Optical Networks 345
5.12.8 Wavelength Planning 346
5.12.9 Transparency 348
Summary 348
Further Reading 348
Problems 349
References 356

II Networks 361

6 Client Layers of the Optical Layer 363
6.1 SONET/SDH . . . 364
6.1.1 Multiplexing . . . 367
6.1.2 SONET/SDH Layers . . . 370
6.1.3 SONET Frame Structure . . . 371
6.1.4 SONET/SDH Physical Layer . . . 375
6.1.5 Elements of a SONET/SDH Infrastructure . . . 378
6.2 ATM . . . 381
6.2.1 Functions of ATM . . . 382
6.2.2 Adaptation Layers . . . 385
6.2.3 Quality of Service . . . 386
6.2.4 Flow Control . . . 387
6.2.5 Signaling and Routing . . . 387
6.3 IP . . . 388
6.3.1 Routing and Forwarding . . . 390
6.3.2 Quality of Service . . . 392
6.3.3 Multiprotocol Label Switching (MPLS) . . . 392
6.3.4 Whither ATM? . . . 394
6.4 Storage-Area Networks . . . 395
6.4.1 ESCON . . . 396
6.4.2 Fibre Channel . . . 397
6.4.3 HIPPI . . . 397
6.5 Gigabit and 10-Gigabit Ethernet . . . 398
Summary . . . 398
Further Reading . . . 399
Problems . . . 399
References . . . 401

7 WDM Network Elements 403
7.1 Optical Line Terminals . . . 406
7.2 Optical Line Amplifiers . . . 408
7.3 Optical Add/Drop Multiplexers . . . 408
7.3.1 OADM Architectures . . . 411
7.3.2 Reconfigurable OADMs . . . 417
7.4 Optical Crossconnects . . . 419
7.4.1 All-Optical OXC Configurations . . . 425
Summary . . . 428
Further Reading . . . 430

Problems 431
References 433

8 **WDM Network Design** 437
8.1 Cost Trade-Offs: A Detailed Ring Network Example 441
8.2 LTD and RWA Problems 448
8.2.1 Lightpath Topology Design 449
8.2.2 Routing and Wavelength Assignment 454
8.2.3 Wavelength Conversion 457
8.2.4 Relationship to Graph Coloring 461
8.3 Dimensioning Wavelength-Routing Networks 462
8.4 Statistical Dimensioning Models 464
8.4.1 First-Passage Model 466
8.4.2 Blocking Model 467
8.5 Maximum Load Dimensioning Models 475
8.5.1 Offline Lightpath Requests 475
8.5.2 Online RWA in Rings 481
Summary 482
Further Reading 484
Problems 484
References 488

9 **Control and Management** 495
9.1 Network Management Functions 495
9.1.1 Management Framework 497
9.1.2 Information Model 499
9.1.3 Management Protocols 500
9.2 Optical Layer Services and Interfacing 502
9.3 Layers within the Optical Layer 504
9.4 Multivendor Interoperability 505
9.5 Performance and Fault Management 507
9.5.1 The Impact of Transparency 508
9.5.2 BER Measurement 509
9.5.3 Optical Trace 509
9.5.4 Alarm Management 510
9.5.5 Data Communication Network (DCN) and Signaling 512
9.5.6 Policing 513
9.5.7 Optical Layer Overhead 514
9.6 Configuration Management 519
9.6.1 Equipment Management 519

9.6.2 Connection Management 520
9.6.3 Adaptation Management 524
9.7 Optical Safety 526
9.7.1 Open Fiber Control Protocol 528
Summary 530
Further Reading 531
Problems 532
References 534

10 Network Survivability 537
10.1 Basic Concepts 539
10.2 Protection in SONET/SDH 542
10.2.1 Point-to-Point Links 544
10.2.2 Self-Healing Rings 546
10.2.3 Unidirectional Path-Switched Rings 549
10.2.4 Bidirectional Line-Switched Rings 550
10.2.5 Ring Interconnection and Dual Homing 555
10.3 Protection in IP Networks 558
10.4 Why Optical Layer Protection 560
10.4.1 Service Classes Based on Protection 566
10.5 Optical Layer Protection Schemes 567
10.5.1 1 + 1 OMS Protection 570
10.5.2 1:1 OMS Protection 571
10.5.3 OMS-DPRing 571
10.5.4 OMS-SPRing 571
10.5.5 1:N Transponder Protection 573
10.5.6 1 + 1 OCh Dedicated Protection 574
10.5.7 OCh-SPRing 575
10.5.8 OCh-Mesh Protection 576
10.5.9 Choice of Protection Technique 582
10.6 Interworking between Layers 582
Summary 583
Further Reading 584
Problems 584
References 587

11 Access Networks 591
11.1 Network Architecture Overview 593
11.2 Enhanced HFC 598
11.3 Fiber to the Curb (FTTC) 599
11.3.1 PON Evolution 609

Summary 610
Further Reading 611
Problems 611
References 612

12 Photonic Packet Switching 615
12.1 Optical Time Division Multiplexing 619
12.1.1 Bit Interleaving 620
12.1.2 Packet Interleaving 623
12.1.3 Optical AND Gates 625
12.2 Synchronization 631
12.2.1 Tunable Delays 632
12.2.2 Optical Phase Lock Loop 633
12.3 Header Processing 634
12.4 Buffering 635
12.4.1 Output Buffering 637
12.4.2 Input Buffering 639
12.4.3 Recirculation Buffering 639
12.4.4 Using Wavelengths for Contention Resolution 641
12.4.5 Deflection Routing 645
12.5 Burst Switching 649
12.6 Testbeds 650
12.6.1 KEOPS 650
12.6.2 NTT's Optical ATM Switches 652
12.6.3 BT Labs Testbeds 654
12.6.4 Princeton University Testbed 654
12.6.5 AON 655
12.6.6 CORD 656
Summary 657
Further Reading 658
Problems 659
References 660

13 Deployment Considerations 667
13.1 The Evolving Telecommunications Network 667
13.1.1 The SONET/SDH Core Network 669
13.1.2 Architectural Choices for Next-Generation Transport Networks 673
13.2 Designing the Transmission Layer 681
13.2.1 Using SDM 682
13.2.2 Using TDM 683
13.2.3 Using WDM 684

13.2.4 Unidirectional versus Bidirectional WDM Systems 686
13.2.5 Long-Haul Networks 688
13.2.6 Long-Haul Network Case Study 689
13.2.7 Long-Haul Undersea Networks 697
13.2.8 Metro Networks 698
13.2.9 Metro Ring Case Study 700
13.2.10 From Opaque Links to Agile All-Optical Networks 704
Summary 705
Further Reading 706
Problems 706
References 710

A Acronyms 711

B Symbols and Parameters 717

C Standards 721
C.1 International Telecommunications Union (ITU-T) 721
C.1.1 Fiber 721
C.1.2 SDH (Synchronous Digital Hierarchy) 721
C.1.3 Optical Networking 722
C.1.4 Management 722
C.2 Telcordia 723
C.2.1 Physical and Environmental 723
C.2.2 SONET 723
C.2.3 Optical Networking 724
C.3 American National Standards Institute (ANSI) 724
C.3.1 SONET 724
C.3.2 ESCON and Fibre Channel 724

D Wave Equations 727

E Pulse Propagation in Optical Fiber 731
E.1 Propagation of Chirped Gaussian Pulses 734
E.2 Nonlinear Effects on Pulse Propagation 735
E.3 Soliton Pulse Propagation 738
Further Reading 739
References 739

F Nonlinear Polarization 741

Ex. 1016.21

G **Multilayer Thin-Film Filters** 743
G.1 Wave Propagation at Dielectric Interfaces 743
G.2 Filter Design 747
References 750

H **Random Variables and Processes** 751
H.1 Random Variables 751
H.1.1 Gaussian Distribution 752
H.1.2 Maxwell Distribution 753
H.1.3 Poisson Distribution 753
H.2 Random Processes 754
H.2.1 Poisson Random Process 755
H.2.2 Gaussian Random Process 756
Further Reading 756
References 756

I **Receiver Noise Statistics** 757
I.1 Shot Noise 759
I.2 Amplifier Noise 760
References 762

Bibliography 763

Index 797

Preface

Since the first edition of this book appeared in February 1998, we have witnessed a dramatic explosion in optical networking. Optical networking used to be confined to a fairly small community of researchers and engineers but is now of great interest to a broad audience including students; engineers in optical component, equipment, and service provider companies; network planners; investors; venture capitalists; and industry and investment analysts.

With the rapid pace in technological advances and the widespread deployment of optical networks over the past three years, the need for a second edition of this book became apparent. In this edition we have attempted to include the latest advances in optical networks and their underlying technologies. We have also tried to make the book more accessible to a broader community of people interested in learning about optical networking. With this in mind, we have rewritten several chapters, added a large amount of new material, and removed some material that is not as relevant to practical optical networks. We have also updated the references and added some new problems.

The major changes we've made are as follows: We have mostly rewritten the introduction to reflect the current understanding of optical networks, and we've added a section called "Transmission Basics" to introduce several terms commonly used in optical networking and wavelength division multiplexing (WDM) to the layperson.

In Chapter 2, we've added significant sections on dispersion management and solitons, along with a section describing the different fiber types now available.

Ex. 1016.23

In Chapter 3, we now cover electro-absorption modulated lasers, tunable lasers, Raman amplifiers, and L-band erbium-doped fiber amplifiers, and we have significantly expanded the section on optical switching to include the new types of switches using micro-electro-mechanical systems (MEMS) and other technologies.

In Chapter 4, we cover return-to-zero modulation and other newer modulation formats such as duobinary, as well as forward error correction, now widely used in high-bit-rate systems. Chapter 5 now includes expanded coverage of chromatic dispersion and polarization effects, which are important factors influencing the design of high-bit-rate long-haul systems.

The networking chapters of the book have been completely rewritten and expanded to reflect the significant progress made in this area. We have organized these chapters as follows: Chapter 6 now includes expanded coverage of SONET/SDH, ATM, and IP networks. Chapter 7 is devoted to architectural considerations underlying WDM network elements. Chapter 8 attempts to provide a unified view of the problems associated with network design and routing in optical networks. Chapter 9 provides significantly expanded coverage of network management and control. We have devoted Chapter 10 to network survivability, with a detailed discussion on optical layer protection. Chapter 11 covers access networks with a focus on emerging passive optical networks (PONs). Chapter 12 provides updated coverage of optical packet-switched networks. Finally, Chapter 13 focuses on deployment considerations and is intended to provide the reader with a broad understanding of how telecommunications networks are evolving. It includes a couple of detailed network planning case studies on a typical long-haul and metro network.

There is currently a great deal of standards activity in this field. We've added an appendix listing the relevant standards. We have also added another appendix listing the acronyms used in the book and moved some of the more advanced material on pulse propagation into an appendix.

While we have mostly added new material, we have also removed some chapters present in the first edition. We have eliminated the chapter on broadcast-and-select networks, as these networks are mostly of academic interest today. Likewise, we also removed the chapter describing optical networking testbeds as they are mostly of historical importance at this point. Interested readers can obtain a copy of these chapters on the Internet at *www.mkp.com/opticalnet2*.

Teaching and Learning from This Book

This book can be used as a textbook for graduate courses in electrical engineering or computer science. Much of the material in this book has been covered in courses taught by us. Chapters 2–5 cover components and transmission technology aspects of optical networking, and Chapters 6–13 deal with the networking aspects. To

Ex. 1016.24

understand the networking issues, students will require a basic undergraduate-level knowledge of communication networks. We have tried to make the transmission-related chapters of the book accessible to networking professionals. For example, components are treated first in a simple qualitative manner from the viewpoint of a network designer, but their principle of operation is then explained in detail. Some prior knowledge of semiconductors and electromagnetics will be helpful in appreciating the detailed treatment in some of the sections.

Readers wishing to obtain a broad understanding of the major aspects of optical networking can read Chapters 1, 6, 7, and 13. Those interested in getting a basic appreciation of the underlying components and transmission technologies can read through Chapters 1–5, skipping the quantitative sections.

The book can be the basis for a graduate course in an electrical engineering or computer science curriculum. A networks-oriented course might emphasize network architectures and control and management, by focusing on Chapters 6–13, and skim over the technology portions of the book. Likewise, a course on optical transmission in an electrical engineering department might instead focus on Chapters 2–5 and omit the remaining chapters. Each chapter is accompanied by a number of problems, and instructors may obtain a solution manual by contacting the publisher at *mkp@mkp.com*.

Acknowledgments

We were fortunate to have an outstanding set of reviewers who made a significant effort in reading through the chapters in detail and providing us with many suggestions to improve the coverage and presentation of material. They have been invaluable in shaping this edition. Specifically, we would like to thank Paul Green, Goff Hill, David Hunter, Rao Lingampalli, Alan McGuire, Shawn O'Donnell, Walter Johnstone, Alan Repech, George Stewart, Suresh Subramaniam, Eric Verillow, and Martin Zirngibl. In addition, we would like to acknowledge Bijan Raahemi, Jim Refi, Krishna Thyagarajan, and Mark R. Wilson who provided inputs and comments on specific topics and pointed out some mistakes in the first edition. Mark R. Wilson was kind enough to provide us with several applications-oriented problems from his class, which we have included in this edition. We would also like to thank Amit Agarwal, Shyam Iyer, Ashutosh Kulshreshtha, and Sarath Kumar for the use of their mesh network design tool, Ashutosh Kulshreshtha for also computing the detailed mesh network design example, Tapan Kumar Nayak for computing the lightpath topology design example, Parthasarathi Palai for simulating the EDFA gain curves, and Rajeev Roy for verifying some of our results. As always, we take responsibility for any errors or omissions and would greatly appreciate hearing from you as you discover them. Please email your comments to *mkp@mkp.com*.

4 chapter

Modulation and Demodulation

OUR GOAL IN THIS CHAPTER is to understand the processes of modulation and demodulation of digital signals. We start by discussing *modulation*, which is the process of converting digital data in electronic form to an optical signal that can be transmitted over the fiber. We then study the *demodulation* process, which is the process of converting the optical signal back into electronic form and extracting the data that was transmitted.

Mainly due to various kinds of noise that get added to the signal in the transmission process, decisions about the transmitted bit (0 or 1) based on the received signal are subject to error. We will derive expressions for the bit error rate introduced by the whole transmission process. Subsequently, we discuss how the bit error rate can be reduced, for the same level of noise (more precisely, signal-to-noise ratio) by the use of forward error-correcting codes. We also discuss clock recovery or synchronization, which is the process of recovering the exact transmission rate at the receiver.

With this background, in the next chapter, we will tackle transmission system engineering, which requires careful attention to a variety of impairments that affect system performance.

4.1 Modulation

The most commonly used modulation scheme in optical communication is *on-off keying* (OOK), which is illustrated in Figure 4.1. In this modulation scheme, a 1 bit is

Ex. 1016.26

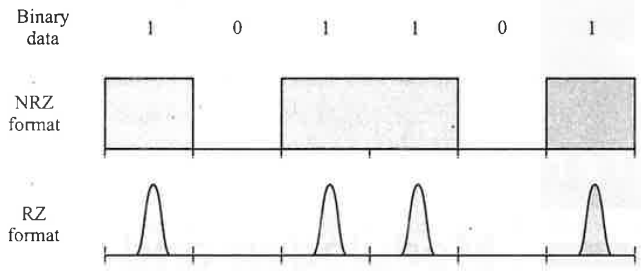

Figure 4.1 On-off keying modulation of binary digital data.

encoded by the presence of a light pulse in the bit interval or by turning a light source (laser or LED) "on." A 0 bit is encoded (ideally) by the absence of a light pulse in the bit interval or by turning a light source "off." The bit interval is the interval of time available for the transmission of a single bit. For example, at a bit rate of 1 Gb/s, the bit interval is 1 ns. As we saw in Section 3.5.4, we can either *directly* modulate the light source by turning it on or off, or use an external modulator in front of the source to perform the same function. Using an external modulator results in less chirp, and thus less of a penalty due to dispersion, and is the preferred approach for high-speed transmission over long distances.

4.1.1 Signal Formats

The OOK modulation scheme can use many different signal formats. The most common signal formats are non-return-to-zero (NRZ) and return-to-zero (RZ). These formats are illustrated in Figure 4.1. In the NRZ format, the pulse for a 1 bit occupies the entire bit interval, and no pulse is used for a 0 bit. If there are two successive 1s, the pulse occupies two successive bit intervals. In the RZ format, the pulse for a 1 bit occupies only a fraction of the bit interval, and no pulse is used for a 0 bit. In electronic (digital) communication, the RZ format has meant that the pulse occupies exactly half the bit period. However, in optical communication, the term RZ is used in a broader sense to describe the use of pulses of duration shorter than the bit period. Thus, there are several variations of the RZ format. In some of them, the pulse occupies a substantial fraction (say, 30%) of the bit interval. The term RZ, without any qualification, usually refers to such systems. If, in addition, the pulses are chirped, they are also sometimes termed dispersion-managed (DM) solitons. In other RZ systems, the pulse occupies only a small fraction of the bit interval. The primary example of such a system is a (conventional) soliton system.

Ex. 1016.27

The major advantage of the NRZ format over the other formats is that the signal occupies a much smaller bandwidth—about half that of the RZ format. The problem with the NRZ format is that long strings of 1s or 0s will result in a total absence of any transitions, making it difficult for the receiver to acquire the bit clock, a problem we discuss in Section 4.4.8. The RZ format ameliorates this problem somewhat since long strings of 1s (but not strings of 0s) will still produce transitions. However, the RZ format requires a higher peak transmit power in order to maintain the same energy per bit, and hence the same bit error rate as the NRZ format.

A problem with all these formats is the lack of *DC balance*. An OOK modulation scheme is said to have DC balance if, for all sequences of data bits that may have to be transmitted, the average transmitted power is constant. It is important for an OOK modulation scheme to achieve DC balance since this makes it easier to set the decision threshold at the receiver (see Section 5.2).

To ensure sufficient transitions in the signal and to provide DC balance, either *line coding* or *scrambling* is used in the system. There are many different types of line codes. One form of a *binary block line code* encodes a block of k data bits into $n > k$ bits that are then modulated and sent over the fiber. At the receiver, the n bits are mapped back into the original k data bits (assuming there were no errors). Line codes can be designed so that the encoded bit sequence is DC balanced and provides sufficient transitions irrespective of the input data bit sequence. An example of such a line code is the (8, 10) code that is used in the Fibre Channel standard [WF83, SV96]. This code has $k = 8$ and $n = 10$. The fiber distributed data interface (FDDI) [Ros86] uses a (4, 5) code that is significantly less complex than this (8, 10) code but does not quite achieve DC balance; the worst-case DC imbalance is 10% [Bur86].

An alternative to using line coding is to use *scrambling*. Scrambling is a one-to-one mapping of the data stream into another data stream before it is transmitted on the link. At the transmitter, a scrambler takes the incoming bits and does an EXOR operation with another carefully chosen sequence of bits. The latter sequence is chosen so as to minimize the likelihood of long sequences of 1s or 0s in the transmitted stream. The data is recovered back at the receiver by a descrambler that extracts the data from the scrambled stream. The advantage of scrambling over line coding is that it does not require any additional bandwidth. The disadvantages are that it does not guarantee DC balance, nor does it guarantee a maximum length for a sequence of 1s or 0s. However, the probability of having long run lengths or DC imbalance is made very small by choosing the mapping so that likely input sequences with long run lengths are mapped into sequences with a small run length. However, since the mapping is one to one, it is possible to choose an input sequence that results in a bad output sequence. The mapping is chosen so that only very rare input sequences produce bad output sequences. See Problem 4.2 for an example of how scrambling is implemented and its properties.

Ex. 1016.28

In practice, the NRZ format is used in most high-speed communication systems, ranging from speeds of 155 Mb/s to 10 Gb/s. Scrambling is widespread and used in most communication equipment ranging from PC modems to high-speed telecommunications links. High-speed computer data links (for example, Fibre Channel, which operates at 800 Mb/s, and Gigabit Ethernet, which operates at 1 Gb/s) use line codes. See Chapter 6 for a discussion of these protocols.

The RZ format is used in certain high-bit-rate communication systems, such as chirped RZ or DM soliton systems (see Section 2.5.1). In these systems, the pulse occupies about half the bit interval, though this is usually not precise as in digital/electronic communication. The use of RZ pulses also minimizes the effects of chromatic dispersion (see Section 5.7.2). RZ modulation with pulses substantially shorter than the bit interval is used in soliton communication systems (see Section 2.5). The pulses need to be very short in such systems because they must be widely separated (by about five times their width) in order to realize the dispersion-free propagation properties of solitons.

4.2 Subcarrier Modulation and Multiplexing

The optical signal emitted by a laser operating in the 1310 or 1550 nm wavelength band has a center frequency around 10^{14} Hz. This frequency is the *optical carrier* frequency. In what we have studied so far, the data modulates this optical carrier. In other words, with an OOK signal, the optical carrier is simply turned on or off, depending on the bit to be transmitted.

Instead of modulating the optical carrier directly, we can have the data first modulate an electrical carrier in the microwave frequency range, typically ranging from 10 MHz to 10 GHz, as shown in Figure 4.2. The upper limit on the carrier frequency is determined by the modulation bandwidth available from the transmitter. The modulated microwave carrier then modulates the optical transmitter. If the transmitter is directly modulated, then changes in the microwave carrier amplitude get reflected as changes in the transmitted optical power envelope, as shown in Figure 4.2. The microwave carrier can itself be modulated in many different ways, including amplitude, phase, and frequency modulation, and both digital and analog modulation techniques can be employed. The figure shows an example where the microwave carrier is amplitude modulated by a binary digital data signal. The microwave carrier is called the *subcarrier*, with the optical carrier being considered the main carrier. This form of modulation is called *subcarrier modulation*.

The main motivation for using subcarrier modulation is to multiplex multiple data streams onto a single optical signal. This can be done by combining multiple microwave carriers at different frequencies and modulating the optical transmitter

Ex. 1016.29

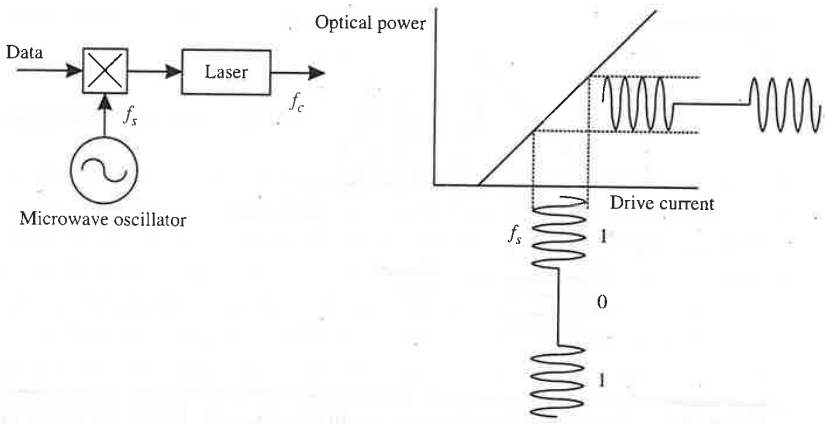

Figure 4.2 Subcarrier modulation. The data stream first modulates a microwave carrier, which, in turn, modulates the optical carrier.

with the combined signal. At the receiver, the signal is detected like any other signal, and the rest of the processing, to separate the subcarriers and extract the data from each subcarrier, is done electronically. This form of multiplexing is called *subcarrier multiplexing* (SCM).

4.2.1 Clipping and Intermodulation Products

The main issue in the design of SCM systems is the trade-off between power efficiency and signal fidelity. Consider Figure 4.2. The SCM system operates around a mean drive current that determines the average optical power. If the mean drive current is increased, for the same SCM waveform, the output optical power is increased. Thus, to keep the output optical power low, the mean drive current must be kept as low as possible. However, the fidelity of the signal depends critically on the linearity of the laser power as a function of the drive current. If f_i, f_j, and f_k denote microwave subcarrier frequencies that are being used, any nonlinearity in laser's power versus drive current characteristic leads to signals at the frequencies $f_i \pm f_j \pm f_k$, which leads to crosstalk, just as in the case of four-wave mixing (see Section 2.4.8). These spurious signals are called *intermodulation products*. Note that the frequencies in the case of SCM are microwave frequencies and those in the FWM case are optical frequencies. But the principle is the same in both cases. For a typical laser, the power–drive current relationship is more linear if the variation in the drive current

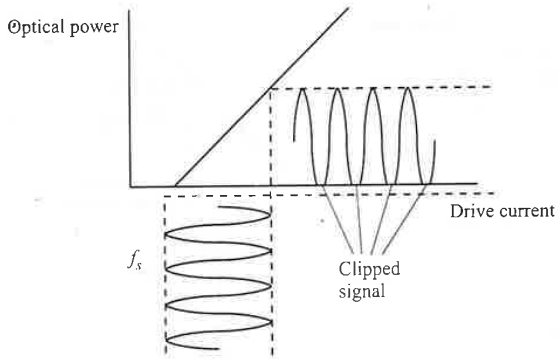

Figure 4.3 Clipping of a subcarrier modulated signal. When the drive current goes below a threshold, the laser output power goes to zero and the signal is said to be clipped.

is a smaller fraction of the average drive current. This means that we must operate at a higher output optical power in order to keep the intermodulation products low. SCM systems use lasers that are specially designed to be highly linear.

The microwave frequencies that are being multiplexed are usually chosen to lie within one octave; that is, if f_L is the lowest frequency and f_H is the highest frequency, these satisfy the condition, $f_H < 2f_L$. When this is the case, all sums and differences of two frequencies—which constitute the second-order intermodulation products—lie either below f_L or above f_H. Thus the second-order intermodulation products produce no crosstalk, and the dominant crosstalk is from third-order intermodulation products, which have much lower power.

A second source of signal distortion in SCM systems is *clipping*. To understand this phenomenon, assume k sinusoids with equal (drive current) amplitude a are being multiplexed. The maximum amplitude of the resulting signal will be ka, and this occurs when all the k signals are in phase. Ideally, the mean operating drive current must be chosen to be greater than ka so that the drive current is nonzero even if all the sinusoids line up in phase. If the operating current is less than ka and all the signals add in phase, there will be no output power for a brief period, when the total current exceeds ka. During this period, the signal is said to be *clipped*. Clipping is illustrated in Figure 4.3 for a single sinusoidal signal.

If k is large, the drive current ka may correspond to a very large optical power. Since the sinusoids are of different frequencies. the probability that they will all add in phase is quite small, particularly for large k. Thus SCM systems are designed to

allow a small clipping probability (a few percent), which substantially reduces the power requirement while introducing only a small amount of signal distortion.

4.2.2 Applications of SCM

SCM is widely used by cable operators today for transmitting multiple analog video signals using a single optical transmitter. SCM is also being used in metropolitan-area networks to combine the signals from various users using electronic FDM followed by SCM. This reduces the cost of the network since each user does not require an optical transmitter/laser. We will study these applications further in Chapter 11.

SCM is also used to combine a control data stream along with the actual data stream. For example, most WDM systems that are deployed carry some control information about each WDM channel along with the data that is being sent. This control information has a low rate and modulates a microwave carrier that lies above the data signal bandwidth. This modulated microwave carrier is called a *pilot tone*. We will discuss the use of pilot tones in Chapter 9.

Often it is necessary to receive the pilot tones from all the WDM channels for monitoring purposes, but not the data. This can be easily done if the pilot tones use different microwave frequencies. If this is the case, and the combined WDM signal is photodetected, the detector output will contain an electronic FDM signal consisting of all the pilot tones from which the control information can be extracted. The information from all the data channels will overlap with one another and be lost.

4.3 Spectral Efficiency

We saw in Chapter 2 that the ultimate bandwidth available in silica optical fiber is about 400 nm from 1.2 μm to 1.6 μm, or about 50 THz. The natural question that arises is, therefore, what is the total capacity at which signals can be transmitted over optical fiber?

There are a few different ways to look at this. The *spectral efficiency* of a digital signal is defined as the ratio of the bit rate to the bandwidth used by the signal. The spectral efficiency depends on the type of modulation and coding scheme used. Today's systems primarily use on-off keying of digital data, and in theory can achieve a spectral efficiency of 1 b/s/Hz. In practice, the spectral efficiency of these systems is more like 0.4 b/s/Hz. Using this number, the maximum capacity of optical fiber is about 20 Tb/s. The spectral efficiency can be improved by using more sophisticated modulation and coding schemes, leading to higher channel capacities than the number above. As spectral efficiency becomes increasingly important, such new schemes are being invented, typically based on proven electrical counterparts.

One such scheme that we discuss in the next section is *optical duobinary modulation.* It can increase the spectral efficiency by a factor of about 1.5, typically, achieving a spectral efficiency of 0.6 b/s/Hz.

4.3.1 Optical Duobinary Modulation

The fundamental idea of duobinary modulation (electrical or optical) is to deliberately introduce intersymbol interference (ISI) by overlapping data from adjacent bits. This is accomplished by adding a data sequence to a 1-bit delayed version of itself. For example, if the (input) data sequence is $(0, 0, 1, 0, 1, 0, 0, 1, 1, 0)$, we would instead transmit the (output) data sequence $(0, 0, 1, 0, 1, 0, 0, 1, 1, 0) + (*, 0, 0, 1, 0, 1, 0, 0, 1, 1) = (0, 0, 1, 1, 1, 1, 0, 1, 2, 1)$. Here the $*$ denotes the initial value of the input sequence, which we assume to be zero.

Note that while the input sequence is binary and consists of 0s and 1s, the output sequence is a ternary sequence consisting of 0s, 1s, and 2s. Mathematically, if we denote the input sequence by $x(nT)$ and the output sequence by $y(nT)$, duobinary modulation results if

$$y(nT) = x(nT) + x(nT - T),$$

where T is the bit period. In the example above, $x(nT) = (0, 0, 1, 0, 1, 0, 0, 1, 1, 0)$, $1 \le n \le 10$, and $y(nT) = (0, 0, 1, 1, 1, 1, 0, 1, 2, 1)$, $1 \le n \le 10$.

Since the bits overlap with each other, how do we recover the input sequence $x(nT)$ at the receiver from $y(nT)$? This can be done by constructing the signal $z(nT) = y(nT) - z(nT - T)$ at the receiver. Note that here we subtract a delayed version of $z(nT)$ from $y(nT)$, and not a delayed version of $y(nT)$ itself. This operation recovers $x(nT)$ since $z(nT) = x(nT)$, assuming we also initialize the sequence $z(0) = 0$. (For readers familiar with digital filters, $y(nT)$ is obtained from $x(nT)$ by a digital filter, and $z(nT)$ from $y(nT)$ by using the inverse of the same digital filter.) The reader should verify this by calculating $z(nT)$ for the example sequence above. To see that this holds generally, just calculate as follows:

$$\begin{aligned}
z(nT) &= y(nT) - z(nT - T) \\
&= y(nT) - y(nT - T) + z(nT - 2T) \\
&= y(nT) - y(nT - T) + y(nT - 2T) - z(nT - 3T) \\
&= y(nT) - y(nT - T) + y(nT - 2T) - \ldots + (-1)^{n-1} y(T) \\
&= [x(nT) + x(nT - T)] - [x(nT - T) - x(nT - 2T)] + \ldots \\
&= x(nT) \qquad (4.1)
\end{aligned}$$

There is one problem with this scheme, however; a single transmission error will cause all further bits to be in error, until another transmission error occurs to correct the first one! This phenomenon is known as *error propagation.* To visualize error propagation, assume a transmission error occurs in some ternary digit in the example sequence $y(nT)$ above, and calculate the decoded sequence $z(nT)$.

The solution to the error propagation problem is to encode the actual data to be transmitted, not by the absolute value of the input sequence $x(nT)$, but by changes in the sequence $x(nT)$. Thus the sequence $x(nT) = (0, 0, 1, 0, 1, 0, 0, 1, 1, 0)$ would correspond to the data sequence $d(nT) = (0, 0, 1, 1, 1, 1, 0, 1, 0, 1)$. A 1 in the sequence $d(nT)$ is encoded by changing the sequence $x(nT)$ from a 0 to a 1, or from a 1 to a 0. To see how differential encoding solves the problem, observe that if a sequence of consecutive bits are all in error, their differences will still be correct, modulo 2.

Transmission of a ternary sequence using optical intensity modulation (the generalization of OOK for nonbinary sequences) will involve transmitting three different optical powers, say, 0, P, and $2P$. Such a modulation scheme will also considerably complicate the demodulation process. We would like to retain the advantage of binary signaling while employing duobinary signaling to reduce the transmission bandwidth.

To see how this can be done, compare $y(nT) = (0, 0, 1, 1, 1, 1, 0, 1, 2, 1)$ and $d(nT) = (0, 0, 1, 1, 1, 1, 0, 1, 0, 1)$ in our example, and observe that $y(nT) \bmod 2 = d(nT)$! This result holds in general, and thus we may think that we could simply map the 2s in $y(nT)$ to 0s and transmit the resulting binary sequence, which could then be detected using the standard scheme. However, such an approach would eliminate the bandwidth advantage of duobinary signaling, as it should, because in such a scheme the differential encoding and the duobinary encoding have done nothing but cancel the effects of each other. The bandwidth advantage of duobinary signaling can only be exploited by using a ternary signaling scheme. A ternary signaling alternative to using three optical power levels is to use a combination of amplitude and phase modulation. Such a scheme is dubbed optical AM-PSK, and most studies of optical duobinary signaling today are based on AM-PSK.

Conceptually, the carrier is a continuous wave signal, a sinusoid, which we can denote by $a\cos(\omega t)$. The three levels of the ternary signal correspond to $-a\cos(\omega t) = a\cos(\omega t + \pi)$, $0 = 0\cos(\omega t)$, and $a\cos(\omega t)$, which we denote by -1, 0, and $+1$, respectively. The actual modulation is usually accomplished using an external modulator in the Mach-Zehnder arrangement (see Sections 3.3.7 and 3.5.4). These are the three signal levels corresponding to 0, 1, and 2, respectively, in $y(nT)$. This modulation scheme is clearly a combination of amplitude and phase modulation; hence the term AM-PSK. The AM-PSK signal retains the bandwidth advantage of duobinary signaling. However, for a direct detection receiver, the signals

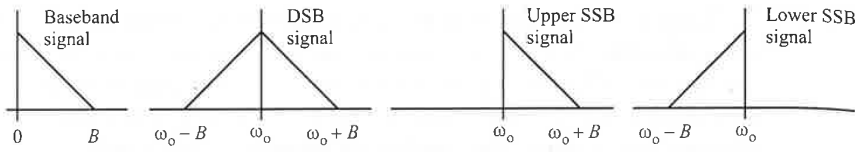

Figure 4.4 Spectrum of a baseband signal compared with the spectra of double sideband (DSB) and single sideband (SSB) modulated signals. The spectral width of the SSB signals is the same as that of the baseband signal, whereas the DSB signal has twice the spectral width of the baseband signal.

$\pm a\cos(\omega t)$ are indistinguishable so that the use of such a receiver merely identifies $2 = 0$ in $y(nT)$ naturally performing the mod 2 operation required to recover $d(nT)$ from $y(nT)$.

4.3.2 Optical Single Sideband Modulation

Another technique for increasing the spectral efficiency is *optical single sideband (SSB) modulation.* Such a scheme can improve the spectral efficiency by a factor of 2, if practical implementations capable of supporting transmission at 10 Gb/s and above can be found. Before we can define what optical SSB modulation is, we need to understand the concept of sidebands in a digital signal.

Consider a sinusoidal carrier signal $\cos(\omega_o t)$. Assume this is directly modulated by a data signal that is also a sinusoid, $\cos(\omega_d t)$, for simplicity. Typically $\omega_d \ll \omega_o$ since ω_o is an optical carrier frequency of the order of 200 THz and ω_d is of the order of 10 GHz. Direct modulation amounts to forming the product $\cos(\omega_o t)\cos(\omega_d t) = 0.5\cos((\omega_o + \omega_d)t) + 0.5\cos((\omega_o - \omega_d)t)$. Thus the transmitted signal contains two sinusoids at $\omega_o + \omega_d$ and $\omega_o + \omega_d$ for a data signal consisting of a single sinusoid at ω_d. In general, for a digital signal with a (baseband) frequency spectrum extending from 0 to B Hz, the modulated signal has a spectrum covering the frequency range from $\omega_o - B$ Hz to $\omega_o + B$ Hz, that is, a range of $2B$ Hz around the carrier frequency ω_o. Each of the spectral bands of width B Hz on either side of the carrier frequency ω_o is called a *sideband,* and such a signal is said to be a *double sideband* (DSB) signal. By appropriate filtering, we can eliminate one of these sidebands: either the lower or the upper one. The resulting signals are called *single sideband* (SSB) signals. DSB and SSB signals are illustrated in Figure 4.4.

The difficulty in implementing optical SSB modulation lies in designing the filters to eliminate one of the sidebands—they have to be very sharp. Instead of filtering it entirely, allowing a small part, or vestige, of one of the sidebands to remain makes

implementation easier. Such a scheme is called *vestigial sideband* (VSB) modulation. This is the modulation scheme used in television systems, and its use is currently being explored for optical systems, mainly for analog signal transmission.

Optical SSB modulation is also being explored today either for analog signal transmission or, equivalently, for SCM systems, which are analog systems from the viewpoint of optical modulation.

4.3.3 Multilevel Modulation

The main technique used in digital communication to achieve spectral efficiencies greater than 1 b/s/Hz is *multilevel modulation.* The simplest multilevel modulation scheme uses $M > 2$ amplitude levels of a sinusoidal carrier to represent M possible signal values. In such a scheme, each signal represents $\log_2 M$ bits. However, the bandwidth occupied by a digital communication system transmitting R such symbols per second is nearly the same as that occupied by an R b/s digital system employing binary signals. Therefore, the bandwidth efficiency of such a multilevel scheme is $\log_2 M$ times higher, and about $\log_2 M$ b/s/Hz. However, such multilevel schemes have not been used in practical optical communication systems to date due to the complexities of detecting such signals at high bit rates. Another potential advantage of multilevel modulation is that the signaling rate on the channel is lower than the data rate. For example, a 16-level modulation scheme would be able to transmit at a date rate of 40 Gb/s but at a signaling rate of 10 Gbaud; that is, each signal occupies a period of 100 ps, and not 25 ps. This, in turn, helps mitigate the effects of dispersion and nonlinearities.

4.3.4 Capacity Limits of Optical Fiber

An upper limit on the spectral efficiency and the channel capacity is given by Shannon's theorem [Sha48]. Shannon's theorem says that the channel capacity C for a binary linear channel with additive noise is given by

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$

Here B is the available bandwidth and S/N is the signal-to-noise ratio. A typical value of S/N is 100. Using this number yields a channel capacity of 350 Tb/s or an equivalent spectral efficiency of 7 b/s/Hz. Clearly, such efficiencies can only be achieved through the use of multilevel modulation schemes.

In practice, today's long-haul systems operate at high power levels to overcome fiber losses and noise introduced by optical amplifiers. At these power levels, nonlinear effects come into play. These nonlinear effects can be thought of as adding

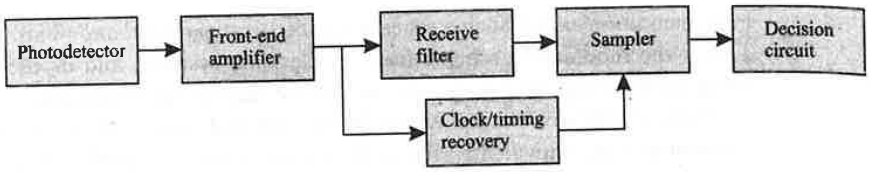

Figure 4.5 Block diagram showing the various functions involved in a receiver.

additional noise, which increases as the transmitted power is increased. Therefore they in turn impose additional limits on channel capacity. Recent work to quantify the spectral efficiency, taking into account mostly cross-phase modulation [Sta99, MS00], shows that the achievable efficiencies are of the order of 3–5 b/s/Hz. Other nonlinearities such as four-wave mixing and Raman scattering may place further limitations. At the same time, we are seeing techniques to reduce the effects of these nonlinearities.

Another way to increase the channel capacity is by reducing the noise level in the system. The noise figure in today's amplifiers is limited primarily by random spontaneous emission, and these are already close to theoretically achievable limits. Advances in quantum mechanics [Gla00] may ultimately succeed in reducing these noise limits.

4.4 Demodulation

The modulated signals are transmitted over the optical fiber where they undergo attenuation and dispersion, have noise added to them from optical amplifiers, and sustain a variety of other impairments that we will discuss in Chapter 5. At the receiver, the transmitted data must be recovered with an acceptable *bit error rate* (BER). The required BER for high-speed optical communication systems today is in the range of 10^{-9} to 10^{-15}, with a typical value of 10^{-12}. A BER of 10^{-12} corresponds to one allowed bit error for every terabit of data transmitted, on average.

Recovering the transmitted data involves a number of steps, which we will discuss in this section. Our focus will be on the demodulation of OOK signals. Figure 4.5 shows the block diagram of a receiver. The optical signal is first converted to an electrical current by a *photodetector*. This electrical current is quite weak and thus we use a *front-end amplifier* to amplify it. The photodetector and front-end amplifier were discussed in Sections 3.6.1 and 3.6.2, respectively.

Ex. 1016.37

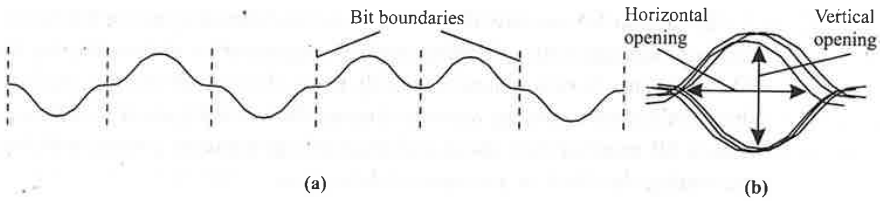

Figure 4.6 Eye diagram. (a) A typical received waveform along with the bit boundaries. (b) The received waveform of (a), wrapped around itself, on the bit boundaries to generate an eye diagram. For clarity, the waveform has been magnified by a factor of 2 relative to (a).

The amplified electrical current is then filtered to minimize the noise outside the bandwidth occupied by the signal. This filter is also designed to suitably shape the pulses so that the bit error rate is minimized. This filter may also incorporate additional functionality, such as minimizing the intersymbol interference due to pulse spreading. If the filter performs this function, it is termed an *equalizer.* The name denotes that the filter equalizes, or cancels, the distortion suffered by the signal. Equalization is discussed in Section 4.4.9.

The signal must then be sampled at the midpoints of the bit intervals to decide whether the transmitted bit in each bit interval was a 1 or a 0. This requires that the bit boundaries be recovered at the receiver. A waveform that is periodic with period equal to the bit interval is called a *clock*. This function is termed *clock recovery,* or *timing recovery,* and is discussed in Section 4.4.8.

A widely used experimental technique to determine the goodness of the received signal is the *eye diagram.* Consider the received waveform shown in Figure 4.6(a). This is a typical shape of the received signal for NRZ modulation, after it has been filtered by the receive filter and is about to be sampled (see Figure 4.5). The bit boundaries are also shown on the figure. If the waveform is cut along at the bit boundaries and the resulting pieces are superimposed on each other, we get the resulting diagram shown in Figure 4.6(b). Such a diagram is called an *eye diagram* because of its resemblance to the shape of the human eye. An eye diagram can be easily generated experimentally using an oscilloscope to display the received signal while it is being triggered by the (recovered) clock. The vertical opening of the eye indicates the margin for bit errors due to noise. The horizontal opening of the eye indicates the margin for timing errors due to an imperfectly recovered clock.

Ex. 1016.38

In Section 1.5, we saw that there could be different types of repeaters, specifically 2R (regeneration with reshaping) and 3R (regeneration with reshaping and retiming). The difference between these primarily lies in the type of receiver used. A 2R receiver does not have the timing recovery circuit shown in Figure 4.5, whereas a 3R does. Also a 3R receiver may use a multirate timing recovery circuit, which is capable of recovering the clock at a variety of data rates.

4.4.1 An Ideal Receiver

In principle, the demodulation process can be quite simple. Ideally, it can be viewed as "photon counting," which is the viewpoint we will take in this section. In practice, there are various impairments that are not accounted for by this model, and we discuss them in the next section.

The receiver looks for the presence or absence of light during a bit interval. If no light is seen, it infers that a 0 bit was transmitted, and if any light is seen, it infers that a 1 bit was transmitted. This is called *direct detection*. Unfortunately, even in the absence of other forms of noise, this will not lead to an ideal error-free system because of the random nature of photon arrivals at the receiver. A light signal arriving with power P can be thought of as a stream of photons arriving at average rate P/hf_c. Here, h is Planck's constant (6.63×10^{-34} J/Hz), f_c is the carrier frequency, and hf_c is the energy of a single photon. This stream can be thought of as a Poisson random process.

Note that our simple receiver does not make any errors when a 0 bit is transmitted. However, when a 1 bit is transmitted, the receiver may decide that a 0 bit was transmitted if no photons were received during that bit interval. If B denotes the bit rate, then the probability that n photons are received during a bit interval $1/B$ is given by

$$e^{-(P/hf_cB)}\frac{\left(\frac{P}{hf_cB}\right)^n}{n!}.$$

Thus the probability of not receiving any photons is $e^{-(P/hf_cB)}$. Assuming equally likely 1s and 0s, the bit error rate of this ideal receiver would be given as

$$\text{BER} = \frac{1}{2}e^{-\frac{P}{hf_cB}}$$

Let $M = P/hf_cB$. The parameter M represents the average number of photons received during a 1 bit. Then the bit error rate can be expressed as

$$\text{BER} = \frac{1}{2}e^{-M}$$

This expression represents the error rate of an ideal receiver and is called the *quantum limit*. To get a bit error rate of 10^{-12}, note that we would need an average of $M = 27$ photons per 1 bit.

In practice, most receivers are not ideal, and their performance is not as good as that of the ideal receiver because they must contend with various other forms of noise, as we shall soon see.

4.4.2 A Practical Direct Detection Receiver

As we have seen in Section 3.6 (see Figure 3.61), the optical signal at the receiver is first photodetected to convert it into an electrical current. The main complication in recovering the transmitted bit is that in addition to the photocurrent due to the signal there are usually three other additional noise currents. The first is the *thermal noise* current due to the random motion of electrons that is always present at any finite temperature. The second is the *shot noise* current due to the random distribution of the electrons generated by the photodetection process even when the input light intensity is constant. The shot noise current, unlike the thermal noise current, is not added to the generated photocurrent but is merely a convenient representation of the variability in the generated photocurrent as a separate component. The third source of noise is the spontaneous emission due to optical amplifiers that may be used between the source and the photodetector. The amplifier noise currents are treated in Section 4.4.5 and Appendix I. In this section, we will consider only the thermal noise and shot noise currents.

The thermal noise current in a resistor R at temperature T can be modeled as a Gaussian random process with zero mean and autocorrelation function $(4k_BT/R)\delta(\tau)$. Here k_B is Boltzmann's constant and has the value 1.38×10^{-23} J/°K, and $\delta(\tau)$ is the Dirac delta function, defined as $\delta(\tau) = 0, \tau \neq 0$ and $\int_{-\infty}^{\infty} \delta(\tau)d\tau = 1$. Thus the noise is white, and in a bandwidth or frequency range B_e, the thermal noise current has the variance

$$\sigma^2_{\text{thermal}} = (4k_BT/R)B_e.$$

This value can be expressed as $I_t^2 B_e$, where I_t is the parameter used to specify the current standard deviation in units of pA/$\sqrt{\text{Hz}}$. Typical values are of the order of 1 pA/$\sqrt{\text{Hz}}$.

The electrical bandwidth of the receiver, B_e, is chosen based on the bit rate of the signal. In practice, B_e varies from $1/2T$ to $1/T$, where T is the bit period. We will also be using the parameter B_o to denote the optical bandwidth seen by the receiver. The optical bandwidth of the receiver itself is very large, but the value of B_o is usually determined by filters placed in the optical path between the transmitter and receiver.

By convention, we will measure B_e in baseband units and B_o in passband units. Therefore, the minimum possible value of $B_o = 2B_e$, to prevent signal distortion.

As we saw in the previous section, the photon arrivals are accurately modeled by a Poisson random process. The photocurrent can thus be modeled as a stream of electronic charge impulses, each generated whenever a photon arrives at the photodetector. For signal powers that are usually encountered in optical communication systems, the photocurrent can be modeled as

$$I = \bar{I} + i_s,$$

where $\bar{I}$ is a constant current, and i_s is a Gaussian random process with mean zero and autocorrelation $\sigma^2_{\text{shot}}\delta(\tau)$. For *pin* diodes, $\sigma^2_{\text{shot}} = 2e\bar{I}$. This is derived in Appendix I. The constant current $\bar{I} = \mathcal{R}P$, where $\mathcal{R}$ is the responsivity of the photodetector, which was discussed in Section 3.6. Here, we are assuming that the dark current, which is the photocurrent that is present in the absence of an input optical signal, is negligible. Thus the shot noise current is also white and in a bandwidth B_e has the variance

$$\sigma^2_{\text{shot}} = 2e\bar{I}B_e. \tag{4.2}$$

If we denote the load resistor of the photodetector by R_L, the total current in this resistor can be written as

$$I = \bar{I} + i_s + i_t,$$

where i_t has the variance $\sigma^2_{\text{thermal}} = (4k_BT/R_L)B_e$. The shot noise and thermal noise currents are assumed to be independent so that, if B_e is the bandwidth of the receiver, this current can be modeled as a Gaussian random process with mean $\bar{I}$ and variance

$$\sigma^2 = \sigma^2_{\text{shot}} + \sigma^2_{\text{thermal}}.$$

Note that both the shot noise and thermal noise variances are proportional to the bandwidth B_e of the receiver. Thus there is a trade-off between the bandwidth of a receiver and its noise performance. A receiver is usually designed so as to have just sufficient bandwidth to accommodate the desired bit rate so that its noise performance is optimized. In most practical direct detection receivers, the variance of the thermal noise component is much larger than the variance of the shot noise and determines the performance of the receiver.

4.4.3 Front-End Amplifier Noise

We saw in Chapter 3 (Figure 3.61) that the photodetector is followed by a front-end amplifier. Components within the front-end amplifier, such as the transistor, also

Ex. 1016.41

contribute to the thermal noise. This noise contribution is usually stated by giving the *noise figure* of the front-end amplifier. The noise figure F_n is the ratio of the input signal-to-noise ratio (SNR_i) to the output signal-to-noise ratio (SNR_o). Equivalently, the noise figure F_n of a front-end amplifier specifies the factor by which the thermal noise present at the input of the amplifier is enhanced at its output. Thus the thermal noise contribution of the receiver has variance

$$\sigma^2_{\text{thermal}} = \frac{4k_B T}{R_L} F_n B_e \tag{4.3}$$

when the front-end amplifier noise contribution is included. Typical values of F_n are 3–5 dB.

4.4.4 APD Noise

As we remarked in Section 3.6.1, the avalanche gain process in APDs has the effect of increasing the noise current at its output. This increased noise contribution arises from the random nature of the avalanche multiplicative gain, $G_m(t)$. This noise contribution is modeled as an increase in the shot noise component at the output of the photodetector. If we denote the responsivity of the APD by $\mathcal{R}_{\text{APD}}$, and the average avalanche multiplication gain by G_m, the average photocurrent is given by $\bar{I} = \mathcal{R}_{\text{APD}} P = G_m \mathcal{R} P$, and the shot noise current at the APD output has variance

$$\sigma^2_{\text{shot}} = 2eG_m^2 F_A(G_m)\mathcal{R}PB_e. \tag{4.4}$$

The quantity $F_A(G_m)$ is called the *excess noise factor* of the APD and is an increasing function of the gain G_m. It is given by

$$F_A(G_m) = k_A G_m + (1 - k_A)(2 - 1/G_m).$$

The quantity k_A is called the ionization coefficient ratio and is a property of the semiconductor material used to make up the APD. It takes values in the range (0–1). The excess noise factor is an increasing function of k_A, and thus it is desirable to keep k_A small. The value of k_A for silicon (which is used at 0.8 μm wavelength) is $\ll 1$, and for InGaAs (which is used at 1.3 and 1.55 μm wavelength bands) is 0.7.

Note that $F_A(1) = 1$, and thus (4.4) also yields the shot noise variance for a *pin* receiver if we set $G_m = 1$.

4.4.5 Optical Preamplifiers

As we have seen in the previous sections, the performance of simple direct detection receivers is limited primarily by thermal noise generated inside the receiver. The

performance can be improved significantly by using an optical (pre)amplifier after the receiver, as shown in Figure 4.7. The amplifier provides added gain to the input signal. Unfortunately, as we saw in Section 3.4.2, the spontaneous emission present in the amplifier appears as noise at its output. The amplified spontaneous (ASE) noise power at the output of the amplifier for each polarization mode is given by

$$P_N = n_{sp} h f_c (G - 1) B_o, \tag{4.5}$$

where n_{sp} is a constant called the spontaneous emission factor, G is the amplifier gain, and B_o is the optical bandwidth. Two fundamental polarization modes are present in a single-mode fiber, as we saw in Chapter 2. Hence the total noise power at the output of the amplifier is $2P_N$.

The value of n_{sp} depends on the level of population inversion within the amplifier. With complete inversion $n_{sp} = 1$, but it is typically higher, around 2–5 for most amplifiers.

For convenience in the discussions to follow, we define

$$P_n = n_{sp} h f_c.$$

To understand the impact of amplifier noise on the detection of the received signal, consider the optical preamplifier system shown in Figure 4.7, used in front of a standard *pin* direct detection receiver. The photodetector produces a current that is proportional to the incident power. The signal current is given by

$$I = \mathcal{R} G P, \tag{4.6}$$

where P is the received optical power.

The photodetector produces a current that is proportional to the optical power. The optical power is proportional to the square of the electric field. Thus the noise field beats against the signal and against itself, giving rise to noise components referred to as the *signal-spontaneous* beat noise and *spontaneous-spontaneous* beat noise, respectively. In addition, shot noise and thermal noise components are also present.

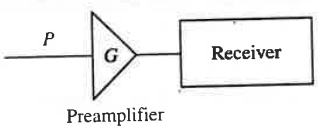

Figure 4.7 A receiver with an optical preamplifier.

The variances of the thermal noise, shot noise, signal-spontaneous noise, and spontaneous-spontaneous noise currents at the receiver are, respectively,

$$\sigma^2_{\text{thermal}} = I_t^2 B_e, \tag{4.7}$$

$$\sigma^2_{\text{shot}} = 2e\mathcal{R}[GP + P_n(G-1)B_o]B_e, \tag{4.8}$$

$$\sigma^2_{\text{sig-spont}} = 4\mathcal{R}^2 G P P_n (G-1) B_e, \tag{4.9}$$

and

$$\sigma^2_{\text{spont-spont}} = 2\mathcal{R}^2[P_n(G-1)]^2(2B_o - B_e)B_e. \tag{4.10}$$

These variances are derived in Appendix I. Here I_t is the receiver thermal noise current. Provided the amplifier gain is reasonably large (> 10 dB), which is usually the case, the shot noise and thermal noise are negligible compared to the signal-spontaneous and spontaneous-spontaneous beat noise. In the bit error rate regime of interest to us (10^{-9} to 10^{-15}), these noise processes can be modeled adequately as Gaussian processes. The spontaneous-spontaneous beat noise can be made very small by reducing the optical bandwidth B_o. This can be done by filtering the amplifier noise before it reaches the receiver. In the limit, B_o can be made as small as $2B_e$. So the dominant noise component is usually signal-spontaneous beat noise.

The amplifier noise is commonly specified by the easily measurable parameter known as the noise figure. Recall from Section 4.4.3 that the noise figure F_n is the ratio of the input signal-to-noise ratio (SNR_i) to the output signal-to-noise ratio (SNR_o). At the amplifier input, assuming that only signal shot noise is present, using (4.2) and (4.6), the SNR is given by

$$\text{SNR}_i = \frac{(\mathcal{R}P)^2}{2\mathcal{R}ePB_e}.$$

At the amplifier output, assuming that the dominant noise term is the signal-spontaneous beat noise, using (4.6) and (4.9), the SNR is given by

$$\text{SNR}_o \approx \frac{(\mathcal{R}GP)^2}{4\mathcal{R}^2 PG(G-1)n_{\text{sp}}hf_cB_e}.$$

The noise figure of the amplifier is then

$$F_n = \frac{\text{SNR}_i}{\text{SNR}_o} \approx 2n_{\text{sp}} \tag{4.11}$$

In the best case, with full population inversion, $n_{\text{sp}} = 1$. Thus the best-case noise figure is 3 dB. Practical amplifiers have a somewhat higher noise figure, typically in

the 4–7 dB range. This derivation assumed that there are no coupling losses between the amplifier and the input and output fibers. Having an input coupling loss degrades the noise figure of the amplifier (see Problem 4.5).

4.4.6 Bit Error Rates

Earlier, we calculated the bit error rate of an ideal direct detection receiver. Next, we will calculate the bit error rate of the practical receivers already considered, which must deal with a variety of different noise impairments.

The receiver makes decisions as to which bit (0 or 1) was transmitted in each bit interval by sampling the photocurrent. Because of the presence of noise currents, the receiver could make a wrong decision resulting in an erroneous bit. In order to compute this bit error rate, we must understand the process by which the receiver makes a decision regarding the transmitted bit.

First, consider a *pin* receiver without an optical preamplifier. For a transmitted 1 bit, let the received optical power $P = P_1$, and let the mean photocurrent $\bar{I} = I_1$. Then $I_1 = \mathcal{R}P_1$, and the variance of the photocurrent is

$$\sigma_1^2 = 2eI_1B_e + 4k_BTB_e/R_L.$$

If P_0 and I_0 are the corresponding quantities for a 0 bit, $I_0 = \mathcal{R}P_0$, and the variance of the photocurrent is

$$\sigma_0^2 = 2eI_0B_e + 4k_BTB_e/R_L.$$

For ideal OOK, P_0 and I_0 are zero, but we will see later (Section 5.3) that this is not always the case in practice.

Let I_1 and I_0 denote the photocurrent sampled by the receiver during a 1 bit and a 0 bit, respectively, and let σ_1^2 and σ_0^2 represent the corresponding noise variances. The noise signals are assumed to be Gaussian. The actual variances will depend on the type of receiver, as we saw earlier. So the bit decision problem faced by the receiver has the following mathematical formulation. The photocurrent for a 1 bit is a sample of a Gaussian random variable with mean I_1 and variance σ_1 (and similarly for the 0 bit as well). The receiver must look at this sample and decide whether the transmitted bit is a 0 or a 1. The possible probability density functions of the sampled photocurrent are sketched in Figure 4.8. There are many possible *decision rules* that the receiver can use; the receiver's objective is to choose the one that minimizes the bit error rate. This *optimum decision rule* can be shown to be the one that, given the observed photocurrent I, chooses the bit (0 or 1) that was *most likely* to have been transmitted. Furthermore, this optimum decision rule can be implemented as follows. Compare the observed photocurrent to a decision threshold I_{th}. If $I \geq I_{\text{th}}$, decide that a 1 bit was transmitted; otherwise, decide that a 0 bit was transmitted.

Ex. 1016.45

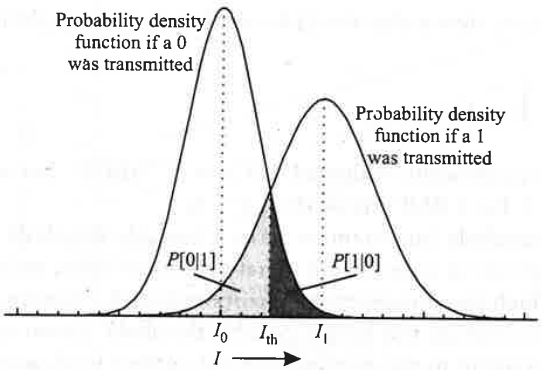

Figure 4.8 Probability density functions for the observed photocurrent.

For the case when 1 and 0 bits are equally likely (which is the only case we consider in this book), the threshold photocurrent is given approximately by

$$I_{th} = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1}. \tag{4.12}$$

This value is very close but not exactly equal to the optimal value of the threshold. The proof of this result is left as an exercise (Problem 4.7). Geometrically, I_{th} is the value of I for which the two densities sketched in Figure 4.8 cross. The probability of error when a 1 was transmitted is the probability that $I < I_{th}$ and is denoted by $P[0|1]$. Likewise, $P[1|0]$ is the probability of deciding that a 1 was transmitted when actually a 0 was transmitted and is the probability that $I \geq I_{th}$. Both probabilities are indicated in Figure 4.8.

Let $Q(x)$ denote the probability that a zero mean, unit variance Gaussian random variable exceeds the value x. Thus

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2}\, dy. \tag{4.13}$$

It now follows that

$$P[0|1] = Q\left(\frac{I_1 - I_{th}}{\sigma_1}\right)$$

and

$$P[1|0] = Q\left(\frac{I_{th} - I_0}{\sigma_0}\right).$$

Using (4.12), it can then be shown that the BER (see Problem 4.6) is given by

$$\text{BER} = Q\left(\frac{I_1 - I_0}{\sigma_0 + \sigma_1}\right). \tag{4.14}$$

The Q function can be numerically evaluated. Let $\gamma = Q^{-1}(\text{BER})$. For a BER rate of 10^{-12}, we need $\gamma \approx 7$. For a BER rate of 10^{-9}, $\gamma \approx 6$.

Note that it is particularly important to have a variable threshold setting in receivers if they must operate in systems with signal-dependent noise, such as optical-amplifier noise. Many high-speed receivers do incorporate such a feature. However, many of the simpler receivers do not have a variable threshold adjustment and set their threshold corresponding to the average received current level, namely, $(I_1 + I_0)/2$. This threshold setting yields a higher bit error rate given by

$$\text{BER} = \frac{1}{2}\left[Q\left(\frac{(I_1 - I_0)}{2\sigma_1}\right) + Q\left(\frac{(I_1 - I_0)}{2\sigma_0}\right)\right].$$

We can use (4.14) to evaluate the BER when the received signal powers for a 0 bit and a 1 bit and the noise statistics are known. Often, we are interested in the inverse problem, namely, determining what it takes to achieve a specified BER. This leads us to the notion of *receiver sensitivity*. The receiver sensitivity $\bar{P}_{\text{sens}}$ is defined as the minimum average optical power necessary to achieve a specified BER, usually 10^{-12} or better. Sometimes the receiver sensitivity is also expressed as the number of photons required per 1 bit, M, which is given by

$$M = \frac{2\bar{P}_{\text{sens}}}{hf_c B},$$

where B is the bit rate.

In the notation introduced earlier, the receiver sensitivity is obtained by solving (4.14) for the average power per bit $(P_0 + P_1)/2$ for the specified BER, say, 10^{-12}. Assuming $P_0 = 0$, this can be obtained as

$$\bar{P}_{\text{sens}} = \frac{(\sigma_0 + \sigma_1)\gamma}{2G_m\mathcal{R}}. \tag{4.15}$$

Here, G_m is the multiplicative gain for APD receivers and is unity for *pin* photodiodes.

First consider an APD or a *pin* receiver, with no optical amplifier in the system. The thermal noise current is independent of the received optical power. However, the shot noise variance is a function of $\bar{P}_{\text{sens}}$. Assume that no power is transmitted for a 0 bit. Then $\sigma_0^2 = \sigma_{\text{thermal}}^2$ and $\sigma_1^2 = \sigma_{\text{thermal}}^2 + \sigma_{\text{shot}}^2$, where the shot noise variance

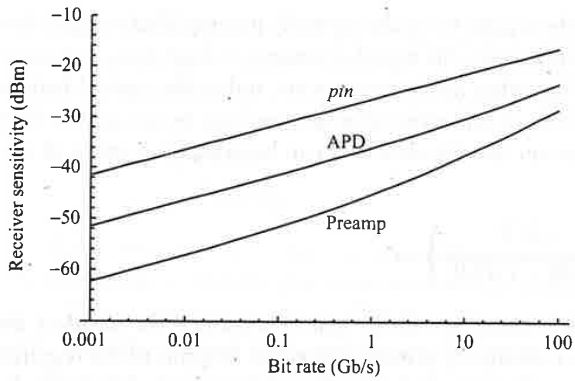

Figure 4.9 Sensitivity plotted as a function of bit rate for typical *pin*, APD, and optically preamplified receivers. The parameters used for the receivers are described in the text.

σ^2_{shot} must be evaluated for the received optical power $P_1 = 2\bar{P}_{\text{sens}}$ that corresponds to a 1 bit. From (4.4),

$$\sigma^2_{\text{shot}} = 4eG_m^2 F_A(G_m)\mathcal{R}\bar{P}_{\text{sens}}B_e.$$

Using this and solving (4.15) for the receiver sensitivity $\bar{P}_{\text{sens}}$, we get

$$\bar{P}_{\text{sens}} = \frac{\gamma}{\mathcal{R}}\left(eB_eF_A(G_m)\gamma + \frac{\sigma_{\text{thermal}}}{G_m}\right). \tag{4.16}$$

Assume that for a bit rate of B b/s, a receiver bandwidth $B_e = B/2$ Hz is required. Let the front-end amplifier noise figure $F_n = 3$ dB and the load resistor $R_L = 100\ \Omega$. Then, assuming the temperature $T = 300°$K, the thermal noise current variance, from (4.3), is

$$\sigma^2_{\text{thermal}} = \frac{4k_BT}{R_L}F_nB_e = 1.656 \times 10^{-22}B\ \text{A}^2. \tag{4.17}$$

Assuming the receiver operates in the 1.55 μm band, the quantum efficiency $\eta = 1$, $\mathcal{R} = 1.55/1.24 = 1.25$ A/W. Using these values, we can compute the sensitivity of a *pin* receiver from (4.16), by setting $G_m = 1$. For BER $= 10^{-12}$ and thus $\gamma \approx 7$, the receiver sensitivity of a *pin* diode is plotted as a function of the bit rate in Figure 4.9. In the same figure, the sensitivity of an APD receiver with $k_A = 0.7$ and an avalanche multiplicative gain $G_m = 10$ is also plotted. It can be seen that the APD receiver has a sensitivity advantage of about 8–10 dB over a *pin* receiver.

We now derive the sensitivity of the optically preamplified receiver shown in Figure 4.7. In amplified systems, the signal-spontaneous beat noise component usually dominates over all the other noise components, unless the optical bandwidth B_o is large, in which case the spontaneous-spontaneous beat noise can also be significant. Making this assumption, the bit error rate can be calculated, using (4.6), (4.9), and (4.14), as

$$\text{BER} = Q\left(\frac{\sqrt{GP}}{2\sqrt{(G-1)P_n B_e}}\right). \tag{4.18}$$

Let us see what receiver sensitivity can be obtained for an ideal preamplified receiver. The receiver sensitivity is measured either in terms of the required power at a particular bit rate or in terms of the number of photons per bit required. As before, we can assume that $B_e = B/2$. Assuming that the amplifier gain G is large and that the spontaneous emission factor $n_{sp} = 1$, we get

$$\text{BER} = Q\left(\sqrt{\frac{M}{2}}\right).$$

To obtain a BER of 10^{-12}, the argument to the $Q(.)$ function γ must be 7. This yields a receiver sensitivity of $M = 98$ photons per 1 bit. In practice, an optical filter is used between the amplifier and the receiver to limit the optical bandwidth B_o and thus reduce the spontaneous-spontaneous and shot noise components in the receiver. For practical preamplified receivers, receiver sensitivities of a few hundred photons per 1 bit are achievable. In contrast, a direct detection pinFET receiver without a preamplifier has a sensitivity of the order of a few thousand photons per 1 bit.

Figure 4.9 also plots the receiver sensitivity for an optically preamplified receiver, assuming a noise figure of 6 dB for the amplifier and an optical bandwidth $B_o =$ 50 GHz that is limited by a filter in front of the amplifier. From Figure 4.9, we see that the sensitivity of a *pin* receiver at a bit rate of 10 Gb/s is −21 dBm and that of an APD receiver is −30 dBm. For 10 Gb/s operation, commercial *pin* receivers with sensitivities of −18 dBm and APD receivers with sensitivities of −24 dBm are available today. From the same figure, at 2.5 Gb/s, the sensitivities of *pin* and APD receivers are −24 dBm and −34 dBm, respectively. Commercial *pin* and APD receivers with nearly these sensitivities at 2.5 Gb/s are available today.

In systems with cascades of optical amplifiers, the notion of sensitivity is not very useful because the signal reaching the receiver already has a lot of added amplifier noise. In this case, the two parameters that are measured are the average received signal power, $\bar{P}_{rec}$, and the received optical noise power, P_{ASE}. The *optical signal-to-noise ratio* (OSNR) is defined as $\bar{P}_{rec}/P_{ASE}$. In the case of an optically preamplified receiver, $P_{ASE} = 2P_n(G-1)B_o$. A system designer needs to relate the

Ex. 1016.49

measured OSNR with the bit error rate. Neglecting the receiver thermal noise and shot noise, it can be shown using (4.6), (4.9), (4.10), and (4.14) that the argument to the $Q(.)$ function, γ, is related to the OSNR as follows:

$$\gamma = \frac{2\sqrt{\frac{B_o}{B_e}}\text{OSNR}}{1+\sqrt{1+4\text{OSNR}}}. \tag{4.19}$$

Consider a typical 2.5 Gb/s system with $B_e = 2$ GHz, with an optical filter with bandwidth $B_o = 36$ GHz placed between the amplifier cascade and the receiver. For $\gamma = 7$, this system requires an OSNR = 4.37, or 6.4 dB. However, this is usually not sufficient because the system must deal with a variety of impairments, such as dispersion and nonlinearities. We will study these in Chapter 5. A rough rule of thumb used by system designers is to design the amplifier cascade to obtain an OSNR of at least 20 dB at the receiver, so as to allow sufficient margin to deal with the other impairments.

4.4.7 Coherent Detection

We saw earlier that simple direct detection receivers are limited by thermal noise and do not achieve the shot noise limited sensitivities of ideal receivers. We saw that the sensitivity could be improved significantly by using an optical preamplifier. Another way to improve the receiver sensitivity is to use a technique called *coherent detection.*

The key idea behind coherent detection is to provide gain to the signal by mixing it with another local light signal from a so-called local-oscillator laser. At the same time, the dominant noise in the receiver becomes the shot noise due to the local oscillator, allowing the receiver to achieve the shot noise limited sensitivity. (In fact, a radio receiver works very much in this fashion except that it operates at radio, rather than light, frequencies.)

A simple coherent receiver is shown in Figure 4.10. The incoming light signal is mixed with a local-oscillator signal via a 3 dB coupler and sent to the photodetector. (We will ignore the 3 dB splitting loss induced by the coupler since it can be eliminated by a slightly different receiver design—see Problem 4.15.) Assume that the phase and polarization of the two waves are perfectly matched. The power seen by the photodetector is then

$$\begin{aligned} P_r(t) &= \left[\sqrt{2aP}\cos(2\pi f_c t) + \sqrt{2P_{\text{LO}}}\cos(2\pi f_{\text{LO}} t)\right]^2 \\ &= aP + P_{\text{LO}} + 2\sqrt{aPP_{\text{LO}}}\cos[2\pi(f_c - f_{\text{LO}})t]. \end{aligned} \tag{4.20}$$

Here, P denotes the input signal power, P_{LO} the local-oscillator power, $a = 1$ or 0 depending on whether a 1 or 0 bit is transmitted (for an OOK signal), and f_c and

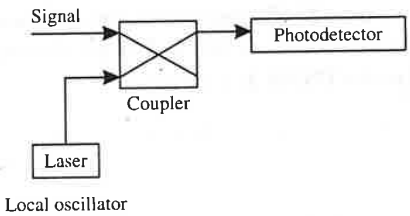

Figure 4.10 A simple coherent receiver.

f_{LO} represent the carrier frequencies of the signal and local-oscillator waves. We have neglected the $2f_c$, $2f_{LO}$, and $f_c + f_{LO}$ components since they will be filtered out by the receiver. In a *homodyne* receiver, $f_c = f_{LO}$, and in a *heterodyne* receiver, $f_c - f_{LO} = f_{IF} \neq 0$. Here, f_{IF} is called the intermediate frequency (IF), typically a few gigahertz.

To illustrate why coherent detection yields improved receiver sensitivities, consider the case of a homodyne receiver. For a 1 bit, we have

$$I_1 = \mathcal{R}(P + P_{LO} + 2\sqrt{PP_{LO}}).$$

and for a 0 bit,

$$I_0 = \mathcal{R}P_{LO}.$$

The key thing to note here is that by making the local-oscillator power P_{LO} sufficiently large, we can make the shot noise dominate over all the other noise components in the receiver. Thus the noise variances are

$$\sigma_1^2 = 2eI_1B_e$$

and

$$\sigma_0^2 = 2eI_0B_e.$$

Usually, P_{LO} is around 0 dBm and P is less than -20 dBm. So we can also neglect P compared to P_{LO} when computing the signal power, and both P and $\sqrt{PP_{LO}}$ compared to P_{LO} when computing the noise variance σ_1^2. With this assumption, using (4.14), the bit error rate is given by

$$\text{BER} = Q\left(\sqrt{\frac{\mathcal{R}P}{2eB_e}}\right).$$

As before, assuming $B_e = B/2$, this expression can be rewritten as

$$\text{BER} = Q(\sqrt{M}),$$

where M is the number of photons per 1 bit as before. For a BER of 10^{-12}, we need the argument of the $Q(.)$ function γ to be 7. This yields a receiver sensitivity of 49 photons per 1 bit, which is significantly better than the sensitivity of a simple direct detection receiver.

However, coherent receivers are generally quite complex to implement and must deal with a variety of impairments. Note that in our derivation we assumed that the phase and polarization of the two waves match perfectly. In practice, this is not the case. If the polarizations are orthogonal, the mixing produces no output. Thus coherent receivers are highly sensitive to variations in the polarizations of the signal and local-oscillator waves as well as any phase noise present in the two signals. There are ways to get around these obstacles by designing more complicated receiver structures [KBW96, Gre93]. However, direct detection receivers with optical preamplifiers, which yield comparable receiver sensitivities, provide a simpler alternative and are widely used today.

There is another advantage to be gained by using coherent receivers in a multichannel WDM system. Instead of using a demultiplexer or filter to select the desired signal optically, with coherent receivers, this selection can be done in the IF domain using electronic filters, which can be designed to have very sharp skirts. This allows very tight channel spacings to be achieved. In addition, in a WDM system, the receiver can be tuned between channels in the IF domain, allowing for rapid tunability between channels, a desirable feature to support fast packet switching. However, we will require highly wavelength-stable and controllable lasers and components to make use of this benefit. Such improvements may result in the resurrection of coherent receivers when WDM systems with large numbers of channels are designed in the future.

4.4.8 Timing Recovery

The process of determining the bit boundaries is called *timing recovery*. The first step is to extract the clock from the received signal. Recall that the clock is a periodic waveform whose period is the bit interval (Section 4.4). This clock is sometimes sent separately by the transmitter, for example, in a different frequency band. Usually, however, the clock must be extracted from the received signal. Even if the extracted clock has a period equal to the bit interval, it may still be out of phase with the received signal; that is, the clock may be offset from the bit boundaries. Usually,

Ex. 1016.52

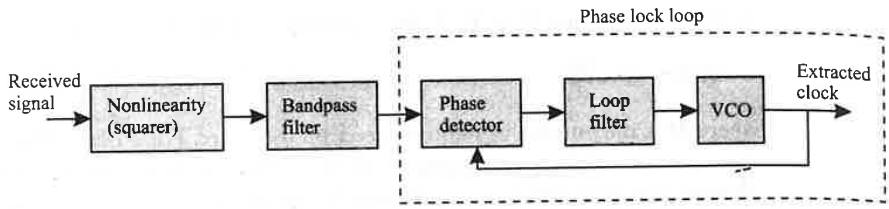

Figure 4.11 Block diagram illustrating timing, or clock, recovery at the receiver.

both the clock frequency (periodicity) and its phase are recovered simultaneously by a single circuit, as shown in Figure 4.11.

If we pass the received signal through a nonlinearity, typically some circuit that calculates the square of the received signal, it can be shown that the result contains a spectral component at $1/T$, where T is the bit period. Thus, we can filter the result using a bandpass filter as shown in Figure 4.11 to get a waveform that is approximately periodic with period T, and which we call a *timing signal.* However, this waveform will still have considerable *jitter;* that is, successive "periods" will have slightly different durations. A "clean" clock with low jitter can be obtained by using the *phase lock loop* (PLL) circuit shown in Figure 4.11.

A PLL consists of a voltage-controlled oscillator (VCO), a phase detector, and a loop filter. A VCO is an oscillator whose output frequency can be controlled by an input voltage. A *phase detector* produces an error signal that depends on the difference in phase between its two inputs. Thus, if the timing signal and the output of the VCO are input to the phase detector, it produces an error signal that is used to adjust the output of the VCO to match the (average) frequency and phase of the timing signal. When this adjustment is complete, the output of the VCO serves as the clock that is used to sample the filtered signal in order to decide upon the values of the transmitted bits. The *loop filter* shown in Figure 4.11 is a critical element of a PLL and determines the residual jitter in the output of the VCO, as well as the ability of the PLL to track changes in the frequency and phase of the timing signal.

4.4.9 Equalization

We remarked in Section 4.4 with reference to Figure 4.5 that the receive filter that is used just prior to sampling the signal can incorporate an *equalization filter* to cancel the effects of intersymbol interference due to pulse spreading. From the viewpoint of the electrical signal that has been received, the entire optical system (including the laser, the fiber, and the photodetector) constitutes the *channel* over which the signal

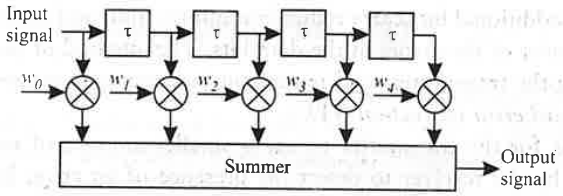

Figure 4.12 A transversal filter, a commonly used structure for equalization. The output (equalized) signal is obtained by adding together suitably delayed versions of the input signal, with appropriate weights.

has been transmitted. If nonlinearities are ignored, the main distortion caused by this channel is the dispersion-induced broadening of the (electrical) pulse. Dispersion is a linear effect, and hence the effect of the channel on the pulse, due to dispersion, can be modeled by the response of a filter with transfer function $H_D(f)$. Hence, in principle, by using the inverse of this filter, say, $H_D^{-1}(f)$, as the equalization filter, this effect can be canceled completely at the receiver. This is what an equalization filter attempts to accomplish.

The effect of an equalization filter is very similar to the effect of dispersion compensating fiber (DCF). The only difference is that in the case of DCF, the equalization is in the optical domain, whereas equalization is done electrically when using an equalization filter. As in the case of DCF, the equalization filter depends not only on the type of fiber used but also on the fiber length.

A commonly used filter structure for equalization is shown in Figure 4.12. This filter structure is called a *transversal filter*. It is essentially a tapped delay line: the signal is delayed by various amounts and added together with individual weights. The choice of the weights, together with the delays, determines the transfer function of the equalization filter. The weights of the tapped delay line have to be adjusted to provide the best possible cancellation of the dispersion-induced pulse broadening.

Electronic equalization involves a significant amount of processing that is difficult to do at higher bit rates, such as 10 Gb/s. Thus optical techniques for dispersion compensation, such as the use of DCF for chromatic dispersion compensation, are currently much more widely used compared to electronic equalization.

4.5 Error Detection and Correction

An error-correcting code is a technique for reducing the bit error rate on a communicaton channel. It involves transmitting additional bits, called *redundancy*, along with

the data bits. These additional bits carry redundant information and are used by the receiver to correct most of the errors in the data bits. This method of reducing the error rate by having the transmitter send redundant bits (using an error-correcting code) is called *forward error correction* (FEC).

An alternative is for the transmitter to use a smaller amount of redundancy, which can be used by the receiver to detect the presence of an error, but there is not sufficient redundancy to identify/correct the errors. This approach is used in telecommunication systems based on SONET and SDH to monitor the bit error rate in the received signal. It is also widely used in data communication systems, where the receiver requests the transmitter to resend the data blocks that are detected to be in error. This technique is called *automatic repeat request* (ARQ).

A simple example of an error-detecting code is the *bit interleaved parity* (BIP) code. A BIP-N code adds N additional bits to the transmitted data. We can use either *even* or *odd* parity. With a BIP-N of even parity, the transmitter computes the code as follows: The first bit of the code provides even parity over the first bit of all N-bit sequences in the covered portion of the signal, the second bit provides even parity over the second bits of all N-bit sequences within the specified portion, and so on. Even parity is generated by setting the BIP-N bits so that there are an even number of 1s in each of all N-bit sequences including the BIP-N bit. Problem 4.16 provides more details on this code.

Optical communication systems are expected to operate at a very low residual BER: 10^{-12} or lower. When the demands on the communication system were lower due to relaxed channel spacing, negligible component crosstalk, negligible effect of nonlinearities, and so on, all that was required to achieve the specified BER was to increase the received power. No FEC techniques were necessary. However, this is not always possible in the very high capacity WDM systems that are being deployed today.

One reason for using FEC instead of higher power is that fiber nonlinearities prevent further increases in transmit power. A second reason for using an FEC code is simply the cost–performance trade-off. The use of an FEC enables a longer communication link before regeneration, since the link can now operate at a lower received power for the same BER. The price to be paid for this is the additional processing involved, mainly at the receiver.

Several communication systems suffer from a BER floor problem: the BER cannot be decreased further by increasing the received power. This is because the main impairment is not due to the various noises (thermal, shot, amplifier) but due to the crosstalk from adjacent WDM channels. Increasing the received power increases the crosstalk proportionately, and thus the BER cannot be decreased beyond a certain

Ex. 1016.55

level, called the *BER floor.* However, FEC can be used to decrease the BER below this floor.

The use of an FEC code can sometimes provide an early warning for BER problems. Assume a link has a BER of 10^{-9} without the use of an FEC. Even though this may be adequate in some situations, it may be better to use an FEC to push the BER down much further, say, to 10^{-15} or lower. Suppose some component fails in such a way as to cause significantly more errors, but does not fail completely. For example, a switch may fail so as to cause significantly more crosstalk, or the output power of a laser may decrease considerably below the specified value. If the system were used without an FEC, the BER may immediately become unacceptable, but with the use of an FEC, the system may be able to continue operation at a much better BER, while alerting the network operator to the problem.

The simplest error-correcting code is a *repetition code.* In such a code, every bit is repeated some number of times, say, thrice. For example, a 1 is transmitted as 111, and a 0 as 000. Thus we have one data, or information bit, plus two redundant bits of the same value. The receiver can estimate the data bit based on the value of the majority of the three received bits. For example, the received bits 101 are interpreted to mean that the data bit is a 1, and the received bits 100 are interpreted to mean the data bit is a 0.

It is easy to see how the use of such a code improves the BER, if the same energy is transmitted per bit after coding, as in the uncoded system. This amounts to transmitting thrice the power in the above example, since three coded bits have to be transmitted for every data bit. In this case, the coded system has the same raw BER—the BER before error correction or decoding—as the uncoded system. However, after decoding, at least two bits in a block of three bits have to be in error for the coded system to make a wrong decision. This substantially decreases the BER of the coded system, as discussed in Problem 4.17. For example, the BER decreases from 10^{-6} for the uncoded system to 3×10^{-12} in the coded system.

However, this is not a fair assessment of the gains due to FEC, since the transmitted power has to be increased by a factor of 3. This may not be possible, for example, if nonlinearities pose a problem, or higher-power lasers are simply unavailable or too expensive. While such a code may have some application in the presence of BER floors, when there are no BER floors, using such a code may defeat the very purpose of using an FEC code. This is because the link length can be increased even further by simply increasing the transmit power and omitting the FEC code. Therefore, a better measure of the performance of an FEC code has to be devised, called the *coding gain.*

The coding gain of an FEC code is the decrease in the receiver sensitivity that it provides for the same BER compared to the uncoded system (for the same transmit

power). In this sense, the repetition code is useless since it has a negative coding gain. However, codes with substantial coding gains, that is, which decrease the BER substantially for the same transmit power as in the uncoded system, have been designed by mathematicians and communication engineers over the last 50 years. In the next section, we discuss a popular and powerful family of such codes called Reed-Solomon codes.

4.5.1 Reed-Solomon Codes

A Reed-Solomon code, named after its inventors Irving Reed and Gus Solomon, does not operate on bits but on groups of bits, which we will call symbols. For example, a symbol could represent a group of 4 bits, or a group of 8 bits (a byte). A transmitter using a Reed-Solomon code considers k data symbols and calculates r additional symbols with redundant information, based on a mathematical formula: the code. The transmitter sends the $n = k + r$ symbols to the receiver.

If the transmitted power is kept constant, since $k+r$ symbols have to be transmitted in the same duration as k symbols, each symbol in the coded system has $k/k+r$ the duration, and hence $k/k+r$ the energy, of a symbol in the uncoded system.

The receiver considers a block of $n = k + r$ symbols, and knowing the code used by the transmitter, it can correctly decode the k data bits even if up to $r/2$ of the $k+r$ symbols are in error.

Reed-Solomon codes have the restriction that if a symbol consists of m bits, the length of the code $n = 2^m - 1$. Thus the code length $n = 255$ if (8-bit) bytes are used as symbols. The number of redundant bits r can take any even value. A popular Reed-Solomon code used in most recently deployed submarine systems has parameters $n = 255$ and $r = 16$, and hence $k = n - r = 239$. In this case, 16 redundant bytes are calculated for every block of 239 data bytes. The number of redundant bits added is less than 7% of the data bits, and the code is capable of correcting up to 8 errored bytes in a block of 239 bytes. This code provides a coding gain of about 6 dB. With this coding gain, the BER can be substantially reduced, for example, from 10^{-5} to 10^{-15}.

A discussion of the encoding and decoding processes involved in the use of Reed-Solomon codes is beyond the scope of this book. A number of references to this topic are listed at the end of this chapter. The principle of operation can be understood based on the following analogy with real numbers.

Assume two real numbers are to be transmitted. Consider a straight line (a polynomial of degree 1), say, $ax + b$, whose two coefficients a and b represent the real numbers to be transmitted. Instead of transmitting a and b, transmit five points on the straight line. The receiver knows that the transmitted points are on a straight line and can recover the straight line, and hence the transmitted data, even if two

Ex. 1016.57

of the five points are in error: it just finds a straight line that fits at least three of the five points. Similarly, if the receiver is given n points but told that they all lie on a degree k polynomial ($k < n$) it can recover the polynomial, even if some of the received points are in error: it just fits the best possible degree-k polynomial to the set of received points.

A Reed-Solomon code works in a similar fashion except that the arithmetic is not over real numbers, but over the finite set of symbols (groups of bits) used in the code. For example, the finite set of symbols consists of the 256 possible 8-bit values for 8-bit symbols. All arithmetic operations are suitably defined over this finite set of symbols, which is called a *finite field*. (If we write 2 = 00000010 and 3 = 00000011, $3/2 \neq 1.5$ in finite field arithmetic: it is some other value in the set of symbols [0, 255].) The $n = 2^m - 1$ transmitted symbols can be viewed as all the possible nonzero values of a degree-k polynomial whose coefficients lie in a finite field of size 2^m. For example, the 255 transmitted values in a Reed-Solomon code with $n = 255$ and $k = 239$ can be viewed as representing the 255 nonzero values of a degree-239 polynomial whose coefficients are 8-bit values that need to be transmitted. The receiver can recover the degree-239 polynomial, and hence the data bits, even if up to 8 of the 255 received values/symbols are in error. (In practice, the data bits are not encoded as the coefficients of such a polynomial, but as the first 239 of the 255 transmitted values/symbols as discussed above.)

FEC is currently used in 10 Gb/s systems and in undersea transmission systems. It is part of the digital wrapper defined by the ITU-T. The two codes standardized by the ITU-T are the (255, 239) and the (255, 223) Reed-Solomon codes. Both are popular codes used in many communication systems, and thus chipsets that implement the encoding and decoding functions for these codes are readily available. The (255, 239) Reed-Solomon code has less than 7% redundancy (16 bytes for 239 bytes) and can correct up to 8 errored bytes in a block of 239 bytes. The (255, 223) Reed-Solomon code has less than 15% redundancy and can correct up to 16 errored bytes in a block of 223 bytes. These codes, as well as much stronger ones, are used today in high-performance optical communication systems.

4.5.2 Interleaving

Frequently, when errors occur, they occur in bursts; that is, a large number of successive bits are in error. The Reed-Solomon codes we studied in the previous section are capable of correcting bursts of errors too. For example, since the (255, 223) code can correct up to 16 errored bytes, it can correct a burst of $16 \times 8 = 128$ bit errors. To correct larger bursts with a Reed-Solomon code, we would have to increase the redundancy. However, the technique of *interleaving* can be used along with the

Ex. 1016.58

Reed-Solomon codes to correct much larger bursts of errors, without increasing the redundancy.

Assume an (n, k) Reed-Solomon code is used and imagine the bytes are arranged in the following order:

1	2	3	...	k	($n - k$ redundant bits)
$k+1$	$k+2$	$k+3$	...	$2k$	($n - k$ redundant bits)
$2k+1$	$2k+2$	$2k+3$	...	$3k$	($n - k$ redundant bits)
...					

Without interleaving, the bytes would be transmitted in row order, that is, the bytes in row 1 are transmitted, followed by the bytes in row 2, and so on.

The idea of interleaving is to transmit the first d bytes in column 1, followed by the first d bytes in column 2, and so on. Thus byte 1 would be followed by byte $k+1$. When d bytes have been transmitted from all n columns, we transmit the next d bytes in column 1 (from rows $(d+1)$ to $2d$), followed by the next d bytes in column 2, and so on. The parameter d is called the *interleaving depth*.

Suppose there is a burst of b byte errors. Only $\lceil b/d \rceil$ of these bytes will occur in the same row due to interleaving. Thus, a (255, 223) Reed-Solomon code will be able to correct any burst of b errors when interleaving to depth d is used, provided $\lceil b/d \rceil < 16$. For example, if interleaving to depth 4 is used ($d = 4$), a (255, 223) Reed-Solomon code can correct a burst of 64 consecutive errored bytes in a block of 223 bytes, though if the errors occur at random byte positions, it can correct only 16 byte errors in the same block size of 223 bytes.

Summary

Modulation is the process of converting data in electronic form to optical form for transmission on the fiber. The simplest form of digital modulation is on-off keying, which most systems use today. Direct modulation of the laser or LED source can be used for transmission at low bit rates over short distances, whereas external modulation is needed for transmission at high bit rates over long distances. Some form of line coding or scrambling is needed to prevent long runs of 1s or 0s in the transmitted data stream to allow the clock to be recovered easily at the receiver and to maintain DC balance.

Subcarrier multiplexing is a technique where many signals are electronically multiplexed using FDM, and the combined signal is used to modulate an optical carrier. Multilevel modulation schemes are more spectrally efficient than on-off keying; optical duobinary signaling is an example of such a scheme.

A simple direct detection receiver looks at the energy received during a bit interval to decide whether it is a 1 or 0 bit. The receiver sensitivity is the average power required at the receiver to achieve a certain bit error rate. The sensitivity of a simple direct detection receiver is determined primarily by the thermal noise in the receiver. The sensitivity can be improved by using APDs instead of *pin* photodetectors or by using an optical preamplifier. Another technique to improve the sensitivity as well as the channel selectivity of the receiver is coherent detection. However, coherent detection is susceptible to a large number of impairments, and it requires a significantly more complicated receiver structure to overcome these impairments. For this reason, it is not practically implemented today.

Clock recovery is an important part of any receiver and is usually based on a phase lock loop.

Electronic equalization is another option to cancel the pulse spreading due to dispersion. This is accomplished by filtering the detected signal electrically to approximately invert the distortion undergone by it.

Error-correcting codes can be used to significantly lower the BER at the expense of additional processing. The most commonly used family of codes are Reed-Solomon codes.

Further Reading

Many books on optical communication cover modulation and detection in greater depth than we have. See, for example, [Gre93, MK88, Agr97]. See also [BL90] for a nice tutorial article on the subject. Subcarrier multiplexing and modulation are treated in depth in [WOS90, OLH89, Dar87, Gre93]. Line coding, scrambling, and bit clock recovery are covered extensively in [LM93]. Optical duobinary modulation is discussed in several recent papers [OY98, Ono98, Fra98]. Optical SSB modulation is discussed in [SNA97, Hui01]. For an excellent and current discussion of channel capacity and information theory in general, we recommend the textbook by Cover and Thomas [CT91]. These techniques have been applied to calculate the capacity limits of optical systems in [MS00].

The principles of signal detection are covered in the classic books by van Trees [vT68] and Wozencraft and Jacobs [WJ90]. For a derivation of shot noise statistics, see [Pap91]. The noise introduced by optical amplifiers has been studied extensively in the literature. Amplifier noise statistics have been derived using quantum mechanical approaches [Per73, Yam80, MYK82, Dan95] as well as semiclassical approaches [Ols89, RH90]. There was a great deal of effort devoted to realizing coherent receivers in the 1980s, but the advent of optical amplifiers in the late 1980s and early 1990s provided a simpler alternative. See [BL90, KBW96] for a detailed treatment

Ex. 1016.60

of coherent receivers. Equalization is treated extensively in many books on digital communication; see, for example, [LM93, Pro00].

The field of error-correcting codes has developed rapidly since its founding by Hamming [Ham50] and Shannon [Sha48] more than a half-century ago. There are many textbooks on this topic; see, for example, [McE77, LC82]. A discussion of FEC techniques in submarine transmission systems appears in [Sab01].

Problems

4.1 A very simple line code used in early data networks is called *bit stuffing*. The objective of this code is to prevent long runs of 1s or 0s but not necessarily achieve DC balance. The encoding works as follows. Suppose the maximum number of consecutive 1s that we are allowed in the bit stream is k. Then the encoder inserts a 0 bit whenever it sees k consecutive 1 bits in the input sequence.

(a) Suppose the incoming data to be transmitted is 111111111001000000 (read left to right). What is the encoded bit stream, assuming $k = 5$?

(b) What is the algorithm used by the decoder to recover the data? Suppose the received bit stream is 011111010111100011 (read left to right). What is the decoded bit stream?

4.2 The SONET standard uses scrambling to prevent long runs of 1s and 0s from occurring in the transmitted bit stream. The scrambling is accomplished by a carefully designed feedback shift register shown in Figure 4.13. The shift register consists of flip-flops whose operation is controlled by a clock running at the bit rate and is reset at the beginning of each frame.

(a) Suppose the incoming data to be transmitted is 11111111111001000000. Assume that the shift register contents are 1111111 at the beginning. What is the scrambled output?

(b) Write a simulation program to compute the scrambled output as a function of the input. The input is a sequence of bits generated by a pseudo-random sequence with equal probabilities for a 1 and a 0. Plot the longest run length of 1s and the longest run length of 0s observed as a function of the sequence length for sequences up to 10 million bits long. Again assume that the shift register contents are 1111111 at the beginning of the sequence. What do you observe?

4.3 Consider the optical duobinary modulation scheme we discussed in Section 4.3.1. If the data sequence is $d(nT) = 10101011010111100001$, calculate (a) the differential

encoding $x(nT)$ of $d(nT)$, and (b) the duobinary encoding $y(nT)$ of $x(nT)$. Recall that $y(nT) \bmod 2 = d(nT)$. How can you compute the sequence $y(nT)$ directly from $d(nT)$ without going through the two-stage differential and duobinary encoding processes?

4.4 Consider the SNR of an APD receiver when both shot noise and thermal noise are present. Assuming that the *excess noise factor* of the APD is given by $F_A(G_m) = G_m^x$ for some $x \in (0, 1)$, derive an expression for the optimum value G_m^{opt} of the APD gain G_m that maximizes the SNR.

4.5 This problem deals with the noise figure of a chain of optical amplifiers and placement of loss elements in the amplifier. The loss element may be an optical add/drop multiplexer, or a gain-flattening filter, or a dispersion compensation module used to compensate for accumulated dispersion along the link. The question is, where should this loss element be placed—in front of the amplifier, after the amplifier, or inside the amplifier?

(a) Consider an optical amplifier with noise figure F. Suppose we introduce a loss element in front of it, with loss $0 < \epsilon \leq 1$ ($\epsilon = 0$ implies no loss, and $\epsilon = 1$ implies 100% loss). Show that the noise figure of the combination is $F/(1 - \epsilon)$. Note that this loss element may also simply reflect the coupling loss into the amplifier. Observe that this combination has a poor noise figure.

(b) Suppose the loss element is placed just after the amplifier. Show that the noise figure of the combination is still F; that is, placing a loss element after the amplifier doesn't affect the noise figure. However, the price we pay in this case is a reduction in optical output power, since the amplifier output is attenuated by the loss element placed after it.

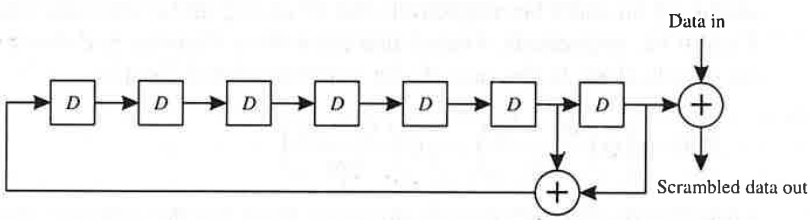

Figure 4.13 The feedback shift register used for scrambling in SONET.

Ex. 1016.62

encoding $x(nT)$ of $d(nT)$, and (b) the duobinary encoding $y(nT)$ of $x(nT)$. Recall that $y(nT) \bmod 2 = d(nT)$. How can you compute the sequence $y(nT)$ directly from $d(nT)$ without going through the two-stage differential and duobinary encoding processes?

4.4 Consider the SNR of an APD receiver when both shot noise and thermal noise are present. Assuming that the *excess noise factor* of the APD is given by $F_A(G_m) = G_m^x$ for some $x \in (0, 1)$, derive an expression for the optimum value G_m^{opt} of the APD gain G_m that maximizes the SNR.

4.5 This problem deals with the noise figure of a chain of optical amplifiers and placement of loss elements in the amplifier. The loss element may be an optical add/drop multiplexer, or a gain-flattening filter, or a dispersion compensation module used to compensate for accumulated dispersion along the link. The question is, where should this loss element be placed—in front of the amplifier, after the amplifier, or inside the amplifier?

(a) Consider an optical amplifier with noise figure F. Suppose we introduce a loss element in front of it, with loss $0 < \epsilon \leq 1$ ($\epsilon = 0$ implies no loss, and $\epsilon = 1$ implies 100% loss). Show that the noise figure of the combination is $F/(1 - \epsilon)$. Note that this loss element may also simply reflect the coupling loss into the amplifier. Observe that this combination has a poor noise figure.

(b) Suppose the loss element is placed just after the amplifier. Show that the noise figure of the combination is still F; that is, placing a loss element after the amplifier doesn't affect the noise figure. However, the price we pay in this case is a reduction in optical output power, since the amplifier output is attenuated by the loss element placed after it.

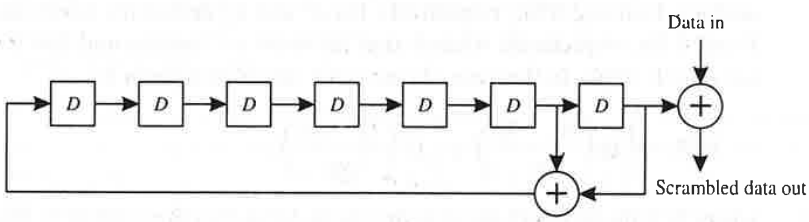

Figure 4.13 The feedback shift register used for scrambling in SONET.

(c) Consider an optical amplifier chain with two amplifiers, with gains G_1 and G_2, respectively, and noise figures F_1 and F_2, respectively, with no loss between the two amplifiers. Assuming $G_1 \gg 1$, show that the noise figure of the combined amplifier chain is

$$F = F_1 + \frac{F_2}{G_1}.$$

In other words, the noise figure of the chain is dominated by the noise figure of the first amplifier, provided its gain is reasonably large, which is usually the case.

(d) Now consider the case where a loss element with loss ϵ is introduced between the first and second amplifier. Assuming $G_1, G_2 \gg 1$, and $(1-\epsilon)G_1G_2 \gg 1$, show that the resulting noise figure of the chain is given by

$$F = F_1 + \frac{F_2}{(1-\epsilon)G_1}.$$

Observe that the loss element doesn't affect the noise figure of the cascade significantly as long as $(1-\epsilon)G_1 \gg 1$, which is usually the case. This is an important fact that is made use of in designing systems. The amplifier is broken down into two stages, the first stage having high gain and a low noise figure, and the loss element is inserted between the two stages. This setup has the advantage that there is no reduction in the noise figure or the output power.

4.6 Show that the BER for an OOK direct detection receiver is given by

$$\text{BER} = Q\left(\frac{I_1 - I_0}{\sigma_0 + \sigma_1}\right).$$

4.7 Consider a binary digital communication system with received signal levels m_1 and m_0 for a 1 bit and 0 bit, respectively. Let σ^2 and σ_0^2 denote the noise variances for a 1 and 0 bit, respectively. Assume that the noise is Gaussian and that a 1 and 0 bit are equally likely. In this case, the bit error rate BER is given by

$$\text{BER} = \frac{1}{2}Q\left(\frac{m_1 - T_d}{\sigma_1}\right) + \frac{1}{2}Q\left(\frac{T_d - m_0}{\sigma_0}\right),$$

where T_d is the receiver's decision threshold. Show that the value of T_d that minimizes the bit error rate is given by

$$T_d = \frac{-m_1\sigma_0^2 + m_0\sigma_1^2 + \sqrt{\sigma_0^2\sigma_1^2(m_1 - m_0)^2 + 2(\sigma_1^2 - \sigma_0^2)\ln(\sigma_1/\sigma_0)}}{\sigma_1^2 - \sigma_0^2}. \tag{4.21}$$

For the case of high signal-to-noise ratios, it is reasonable to assume that

$$(m_1 - m_0)^2 \gg \frac{2(\sigma_1^2 - \sigma_0^2)\ln(\sigma_1/\sigma_0)}{\sigma_0^2\sigma_1^2}.$$

In this case, (4.21) can be simplified to

$$T_d = \frac{m_0\sigma_1 + m_1\sigma_0}{\sigma_1 + \sigma_0}.$$

With $m_1 = RP_1$ and $m_0 = RP_0$, this is the same as (4.12).

4.8 Consider a *pin* direct detection receiver where the thermal noise is the main noise component, and its variance has the value given by (4.17). What is the receiver sensitivity expressed in photons per 1 bit at a bit rate of 100 Mb/s and 1 Gb/s for a bit error rate of 10^{-12}? Assume that the operating wavelength is 1.55 μm and the responsivity is 1.25 A/W.

4.9 Consider the receiver sensitivity, $\bar{P}_{\text{rec}}$ (for an arbitrary BER, not necessarily 10^{-9}), of an APD receiver when both shot noise and thermal noise are present but neglecting the dark current, for direct detection of on-off–keyed signals. Assume no power is transmitted for a 0 bit.

(a) Derive an expression for $\bar{P}_{\text{rec}}$.
(b) Find the optimum value G_m^{opt} of the APD gain G_m that minimizes $\bar{P}_{\text{rec}}$.
(c) For $G_m = G_m^{\text{opt}}$, what is the (minimum) value of $\bar{P}_{\text{rec}}$?

4.10 Derive (4.18).

4.11 Plot the receiver sensitivity as a function of bit rate for an optically preamplified receiver for three different optical bandwidths: (a) the ideal case, $B_o = 2B_e$, (b) $B_o = 100$ GHz, and (c) $B_o = 30$ THz, that is, an unfiltered receiver. Assume an amplifier noise figure of 6 dB, and the electrical bandwidth B_e is half the bit rate, and use the thermal noise variance given by (4.17). What do you observe as the optical bandwidth is increased?

4.12 You are doing an experiment to measure the BER of an optically preamplified receiver. The setup consists of an optical amplifier followed by a variable attenuator to adjust the power going into the receiver, followed by a *pin* receiver. You plot the BER versus the power going into the receiver over a wide range of received powers. Calculate and plot this function. What do you observe regarding the slope of this curve? Assume that $B_o = 100$ GHz, $B_e = 2$ GHz, $B = 2.5$ Gb/s, a noise figure of 6 dB for the optical amplifier, and a noise figure of 3 dB for the front-end amplifier.

4.13 Derive (4.19).

Ex. 1016.65

4.14 Another form of digital modulation that can be used in conjunction with coherent reception is *phase-shift keying* (PSK). Here $\sqrt{2P}\cos(2\pi f_c t)$ is received for a 1 bit and $-\sqrt{2P}\cos(2\pi f_c t)$ is received for a 0 bit. Derive an expression for the bit error rate of a PSK homodyne coherent receiver. How many photons per bit are required to obtain a bit error rate of 10^{-9}?

4.15 A balanced coherent receiver is shown in Figure 4.14. The input signal and local oscillator are sent through a 3 dB coupler, and each output of the coupler is connected to a photodetector. This 3 dB coupler is different in that it introduces an additional phase shift of $\pi/2$ at its second input and second output. The detected current is the difference between the currents generated by the two photodetectors. Show that this receiver structure avoids the 3 dB penalty associated with the receiver we discussed in Section 4.4.7. Use the transfer function for a 3 dB coupler given by (3.1).

4.16 SONET and SDH systems use an 8-bit interleaved parity (BIP-8) check code with even parity to detect errors. The code works as follows. Let $b_0, b_1, b_2, \ldots$ denote the sequence of bits to be transmitted. The transmitter adds an 8-bit code sequence $c_0, c_1, \ldots, c_7$, to the end of this sequence where

$$c_i = b_i \oplus b_{i+8} \oplus b_{i+16} + \cdots.$$

Here $\oplus$ denotes an "exclusive OR" operation ($0 \oplus 0 = 0$, $0 \oplus 1 = 1$, $1 \oplus 1 = 0$).

(a) Suppose the bits to be transmitted are 010111010111101111001110. What is the transmitted sequence with the additional parity check bits?

(b) Suppose the received sequence (including the parity check bits at the end) is 010111010111101111001110. How many bits are in error? Assume that if a parity check indicates an error, it is caused by a single bit error in one of the bits over which the parity is computed.

4.17 If the BER of an uncoded system is p, show that the same system has a BER of $3p^2 + p^3$ when the repetition code (each bit is repeated thrice) is used. Note that the receiver makes its decision on the value of the transmitted bit by taking a majority vote on the corresponding three received coded bits. Assume that the energy per bit remains the same in both cases.

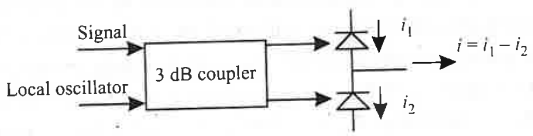

Figure 4.14 A balanced coherent receiver.

References

[Agr97] G. P. Agrawal. *Fiber-Optic Communication Systems*. John Wiley, New York, 1997.

[BL90] J. R. Barry and E. A. Lee. Performance of coherent optical receivers. *Proceedings of IEEE*, 78(8):1369–1394, Aug. 1990.

[Bur86] W. E. Burr. The FDDI optical data link. *IEEE Communications Magazine*, 24(5):18–23, May 1986.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.

[Dan95] S. L. Danielsen et al. Detailed noise statistics for an optically preamplified direct detection receiver. *IEEE/OSA Journal on Lightwave Technology*, 13(5):977–981, 1995.

[Dar87] T. E. Darcie. Subcarrier multiplexing for multiple-access lightwave networks. *IEEE/OSA Journal on Lightwave Technology*, LT-5:1103–1110, 1987.

[Fra98] T. Franck et al. Duobinary transmitter with low intersymbol interference. *IEEE Photonics Technology Letters*, 10:597–599, 1998.

[Gla00] A. M. Glass et al. Advances in fiber optics. *Bell Labs Technical Journal*, 5(1):168–187, Jan.–March 2000.

[Gre93] P. E. Green. *Fiber-Optic Networks*. Prentice Hall, Englewood Cliffs, NJ, 1993.

[Ham50] R. W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 29, 1950.

[Hui01] R. Hui et al. 10 Gb/s SCM system using optical single side-band modulation. In *OFC 2001 Technical Digest*, pages MM4/1–4, 2001.

[KBW96] L. G. Kazovsky, S. Benedetto, and A. E. Willner. *Optical Fiber Communication Systems*. Artech House, Boston, 1996.

[LC82] S. Lin and D. J. Costello. *Error Correcting Codes*. Prentice Hall, Englewood Cliffs, NJ, 1982.

[LM93] E. A. Lee and D. G. Messerschmitt. *Digital Communication*, 2nd edition. Kluwer, Boston, 1993.

[McE77] R. J. McEliece. *The Theory of Information and Coding: A Mathematical Framework for Communication*. Addison-Wesley, Reading, MA, 1977.

[MK88] S D. Miller and I. P. Kaminow, editors. *Optical Fiber Telecommunications II*. Academic Press, San Diego, CA, 1988.

[MS00] P. P. Mitra and J. B. Stark. Nonlinear limits to the information capacity of optical fibre communications. *Nature*, pages 1027–1030, 2000.

[MYK82] T. Mukai, Y. Yamamoto, and T. Kimura. S/N and error-rate performance of AlGaAs semiconductor laser preamplifier and linear repeater systems. *IEEE Transactions on Microwave Theory and Techniques*, 30(10):1548–1554, 1982.

[OLH89] R. Olshanksy, V. A. Lanzisera, and P. M. Hill. Subcarrier multiplexed lightwave systems for broadband distribution. *IEEE/OSA Journal on Lightwave Technology*, 7(9):1329–1342, Sept. 1989.

[Ols89] N. A. Olsson. Lightwave systems with optical amplifiers. *IEEE/OSA Journal on Lightwave Technology*, 7(7):1071–1082, July 1989.

[Ono98] T. Ono et al. Characteristics of optical duobinary signals in terabit/s capacity, high spectral efficiency WDM systems. *IEEE/OSA Journal on Lightwave Technology*, 16:788–797, 1998.

[OY98] T. Ono and Y. Yano. Key technologies for terabit/second WDM systems high spectral efficiency of over 1 bit/s/hz. *IEEE Journal of Quantum Electronics*, 34:2080–2088, 1998.

[Pap91] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*, 3rd edition. McGraw-Hill, New York, 1991.

[Per73] S. D. Personick. Applications for quantum amplifiers in simple digital optical communication systems. *Bell System Technical Journal*, 52(1):117–133, Jan. 1973.

[Pro00] J. G. Proakis. *Digital Communications*, 4th edition. McGraw-Hill, New York, 2000.

[RH90] R. Ramaswami and P. A. Humblet. Amplifier induced crosstalk in multi-channel optical networks. *IEEE/OSA Journal on Lightwave Technology*, 8(12):1882–1896, Dec. 1990.

[Ros86] F. E. Ross. FDDI—a tutorial. *IEEE Communications Magazine*, 24(5):10–17, May 1986.

[Sab01] O. A. Sab. FEC techniques in submarine transmission systems. In *OFC 2001 Technical Digest*, pages TuF1/1–3, 2001.

[Sha48] C E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, July 1948.

[SNA97] G. H. Smith, D. Novak, and Z. Ahmed. Technique for optical SSB generation to overcome dispersion penalties in fibre-radio systems. *Electronics Letters*, 33:74–75, 1997.

[Sta99] J. B. Stark. Fundamental limits of information capacity for optical communications channels. In *Proceedings of European Conference on Optical Communication*, pages 1–28, Nice, France, Sept. 1999.

Ex. 1016.68

[SV96] M. W. Sachs and A. Varma. Fibre channel and related standards. *IEEE Communications Magazine*, 34(8):40–49, Aug. 1996.

[vT68] H. L. van Trees. *Detection, Estimation, and Modulation Theory, Part I.* John Wiley, New York, 1968.

[WF83] A. X. Widmer and P. A. Franaszek. A DC-balanced, partitioned-block, 8B-10B transmission code. *IBM Journal of Research and Development*, 27(5):440–451, Sept. 1983.

[WJ90] J. M. Wozencraft and I. M. Jacobs. *Principles of Communication Engineering.* Waveland Press, Prospect Heights, IL, 1990. Reprint of the originial 1965 edition.

[WOS90] W. I. Way, R. Olshansky, and K. Sato, editors. Special issue on applications of RF and microwave subcarriers to optical fiber transmission in present and future broadband networks. *IEEE Journal of Selected Areas in Communications*, 8(7), Sept. 1990.

[Yam80] Y. Yamamoto. Noise and error-rate performance of semiconductor laser amplifiers in PCM-IM transmission systems. *IEEE Journal of Quantum Electronics*, 16:1073–1081, 1980.

7 chapter

WDM Network Elements

WE HAVE ALREADY EXPLORED some of the motivations for deploying WDM networks in Chapter 1 and will go back to this issue in Chapter 13. These networks provide circuit-switched end-to-end *optical channels*, or *lightpaths*, between network nodes to their users, or *clients*. A lightpath consists of an optical channel, or wavelength, between two network nodes that is routed through multiple intermediate nodes. Intermediate nodes may switch and convert wavelengths. These networks may thus be thought of as *wavelength-routing* networks. Lightpaths are set up and taken down as dictated by the users of the network.

In this chapter we will explore the architectural aspects of the network elements that are part of this network. The architecture of such a network is shown in Figure 7.1. The network consists of *optical line terminals* (OLTs), *optical add/drop multiplexers* (OADMs), and *optical crossconnects* (OXCs) interconnected via fiber links. Not shown in the figure are optical line amplifiers, which are deployed along the fiber link at periodic locations to amplify the light signal. In addition, the OLTs, OADMs, and OXCs may themselves incorporate optical amplifiers to make up for losses. As of this writing, OLTs are widely deployed, and OADMs are deployed to a lesser extent. OXCs are just beginning to be deployed.

The architecture supports a variety of topologies, including ring and mesh topologies. OLTs multiplex multiple wavelengths into a single fiber and also demultiplex a composite WDM signal into individual wavelengths. OLTs are used at either end of a point-to-point link. OADMs are used at locations where some fraction of the wavelengths need to be terminated locally and others need to be routed to other

Ex. 1016.70

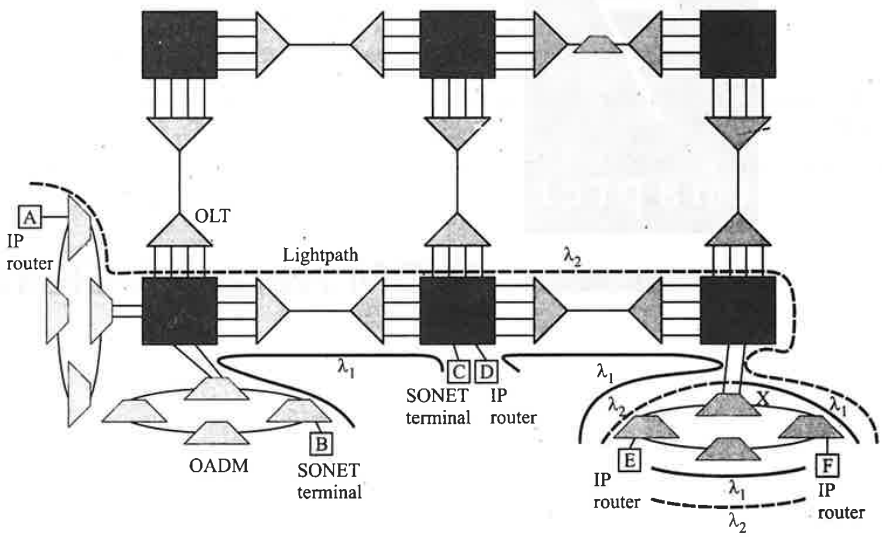

Figure 7.1 A wavelength-routing mesh network showing optical line terminals (OLTs), optical add/drop multiplexers (OADMs), and optical crossconnects (OXCs). The network provides lightpaths to its users, such as SONET boxes and IP routers. A lightpath is carried on a wavelength between its source and destination but may get converted from one wavelength to another along the way.

destinations. They are typically deployed in linear or ring topologies. OXCs perform a similar function but on a much larger scale in terms of number of ports and wavelengths involved, and are deployed in mesh topologies or in order to interconnect multiple rings. We will study these network elements in detail later in this chapter. The users (or clients) of this network are connected to the OLTs, OADMs, or OXCs. The network supports a variety of client types, such as IP routers, ATM switches, and SONET terminals and ADMs.

Each link can support a certain number of wavelengths. The number of wavelengths that can be supported depends on the component- and transmission-imposed limitations that we studied in Chapters 2, 3, and 5.

We next describe several noteworthy features of this architecture:

Wavelength reuse. Observe from Figure 7.1 that multiple lightpaths in the network can use the same wavelength, as long as they do not overlap on any link. This spatial reuse capability allows the network to support a large number of lightpaths using a limited number of wavelengths.

Wavelength conversion. Lightpaths may undergo *wavelength conversion* along their route. Figure 7.1 shows one such lightpath that uses wavelength λ_2 on link EX, gets converted to λ_1 at node X, and uses that wavelength on link XF. Wavelength conversion can improve the utilization of wavelengths inside the network. We will study this aspect in Section 7.4.1 and in Chapter 8. Wavelength conversion is also needed at the boundaries of the network to adapt signals from outside the network into a suitable wavelength for use inside the network.

Transparency. Transparency refers to the fact that the lightpaths can carry data at a variety of bit rates, protocols, and so forth and can, in effect, be made protocol insensitive. This enables the optical layer to support a variety of higher layers *concurrently*. For example, Figure 7.1 shows lightpaths between pairs of SONET terminals, as well as between pairs of IP routers. These lightpaths could carry data at different bit rates and protocols.

Circuit switching. The lightpaths provided by the optical layer can be set up and taken down upon demand. These are analogous to setting up and taking down circuits in circuit-switched networks, except that the rate at which the setup and take-down actions occur is likely to be much slower than, say, the rate for telephone networks with voice circuits. In fact, today these lightpaths, once set up, remain in the network for months to years. With the advent of new services and capabilities offered by today's network equipment, we are likely to see a situation where this process is more dynamic, both in terms of arrivals of lightpath requests and durations of lightpaths.

Note that packet switching is *not* provided within the optical layer. The technology for optical packet switching is still fairly immature; see Chapter 12 for details. It is left to the higher layer, for example, IP or ATM, to perform any packet-switching functions needed.

Survivability. The network can be configured such that, in the event of failures, lightpaths can be rerouted over alternative paths automatically. This provides a high degree of resilience in the network. We will study this aspect further in Chapter 10.

Lightpath topology. The *lightpath topology* is the graph consisting of the network nodes, with an edge between two nodes if there is a lightpath between them. The lightpath topology thus refers to the topology seen by the higher layers using the

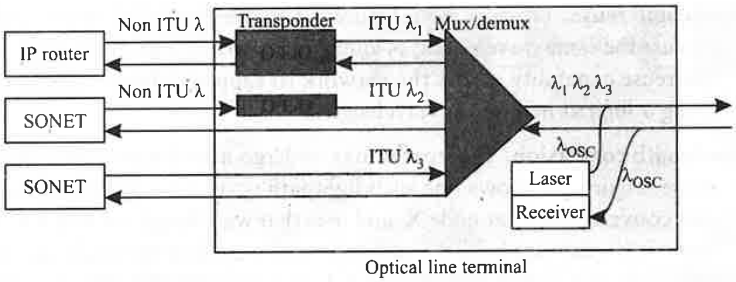

Figure 7.2 Block diagram of an optical line terminal. The OLT has wavelength multiplexers and demultiplexers and adaptation devices called transponders. The transponders convert the incoming signal from the client to a signal suitable for transmission over the WDM link and an incoming signal from the WDM link to a suitable signal toward the client. Transponders are not needed if the client equipment can directly send and receive signals compatible with the WDM link. The OLT also terminates a separate optical supervisory channel (OSC) used on the fiber link.

optical layer. To an IP network residing above the optical layer, the lightpaths look like links between IP routers. The set of lightpaths can be tailored to meet the traffic requirements of the higher layers. This topic will be explored further in Chapter 8.

7.1 Optical Line Terminals

OLTs are relatively simple network elements from an architectural perspective. They are used at either end of a point-to-point link to multiplex and demultiplex wavelengths. Figure 7.2 shows the three functional elements inside an OLT: *transponders, wavelength multiplexers,* and optionally, *optical amplifiers* (not shown in the figure). A transponder adapts the signal coming in from a client of the optical network into a signal suitable for use inside the optical network. Likewise, in the reverse direction, it adapts the signal from the optical network into a signal suitable for the client. The interface between the client and the transponder may vary depending on the client, bit rate, and the distance and/or loss between the client and the transponder. The most common interface is the SONET/SDH short-reach (SR) interface described in Section 6.1.4. We are also seeing the emergence of cheaper very short reach (VSR) interfaces at bit rates of 10 Gb/s and higher.

Ex. 1016.73

The adaptation includes several functions, which we will explore in detail in Section 9.6.3. The signal may need to be converted into a wavelength that is suited for use inside the optical network. The wavelengths generated by the transponder typically conform to standards set by the International Telecommunications Union (ITU) in the 1.55 μm wavelength window, as indicated in the figure, while the incoming signal may be a 1.3 μm signal. The transponder may add additional overhead for purposes of network management. It may also add forward error correction (FEC), particularly for signals at 10 Gb/s and higher rates. The transponder typically also monitors the bit error rate of the signal at the ingress and egress points in the network. For these reasons, the adaptation is typically done through an optical-to-electrical-to-optical (O/E/O) conversion. Down the road, we may see some of the all-optical wavelength-converting technologies of Section 3.8 being used in transponders—these are still in research laboratories.

In some situations, it is possible to have the adaptation enabled only in the incoming direction and have the ITU wavelength in the other direction directly sent to the client equipment. This is shown in the middle of Figure 7.2. In some other situations, we can avoid the use of transponders by having the adaptation function performed inside the client equipment that is using the optical network, such as a SONET network element. This is shown at the bottom of Figure 7.2. This reduces the cost and results in a more compact and power-efficient solution. However, this WDM interface specification is proprietary to each WDM vendor, and there are no standards. (More on this in Section 9.4.) Transponders typically constitute the bulk of the cost, footprint, and power consumption in an OLT. Therefore reducing the number of transponders helps minimize both the cost and the size of the equipment deployed.

The signal coming out of a transponder is multiplexed with other signals at different wavelengths using a wavelength multiplexer onto a fiber. Any of the multiplexing technologies described in Chapter 3, such as arrayed waveguide gratings, dielectric thin-film filters, or fiber Bragg gratings, can be used for this purpose. In addition, an optical amplifier may be used to boost the signal power if needed. In the other direction, the WDM signal is amplified again, if needed, before it is sent through a demultiplexer that extracts the individual wavelengths. These wavelengths are again terminated in a transponder (if present) or directly in the client equipment.

Finally, the OLT also terminates an *optical supervisory channel* (OSC). The OSC is carried on a separate wavelength, different from the wavelengths carrying the actual traffic. It is used to monitor the performance of amplifiers along the link as well as for a variety of other management functions that we will study in Chapter 9.

Ex. 1016.74

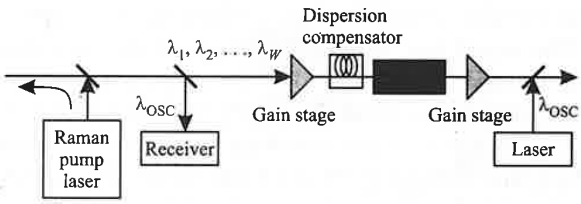

Figure 7.3 Block diagram of a typical optical line amplifier. Only one direction is shown. The amplifier uses multiple erbium gain stages and optionally includes dispersion compensators and OADMs between the gain stages. A Raman pump may be used to provide additional Raman gain over the fiber span. The OSC is filtered at the input and terminated, and added back at the output.

7.2 Optical Line Amplifiers

Optical line amplifiers are deployed in the middle of the optical fiber link at periodic intervals, typically 80–120 km. Figure 7.3 shows a block diagram of a fairly standard optical line amplifier. The basic element is an erbium-doped fiber gain block, which we studied in Chapter 3. Typical amplifiers use two or more gain blocks in cascade, with so-called midstage access. This feature allows some lossy elements to be placed between the two amplifier stages without significantly impacting the overall noise figure of the amplifier (see Problem 4.5 in Chapter 4). These elements include dispersion compensators to compensate for the chromatic dispersion accumulated along the link, and also the OADMs that we will discuss next. The amplifiers also include automatic gain control (see Chapter 5) and built-in performance monitoring of the signal, a topic we will discuss in Chapter 9.

We are also seeing the use of Raman amplifiers, where a high-power pump laser is used at each amplifier site to pump the fiber in the direction opposite to the signal. The optical supervisory channel is filtered at the input and terminated, and added back at the output. In a system using C- and L-bands, the bands are separated at the input to the amplifier and separate EDFAs are used for each band.

7.3 Optical Add/Drop Multiplexers

Optical add/drop multiplexers (OADMs) provide a cost-effective means for handling passthrough traffic in both metro and long-haul networks. OADMs may be used at

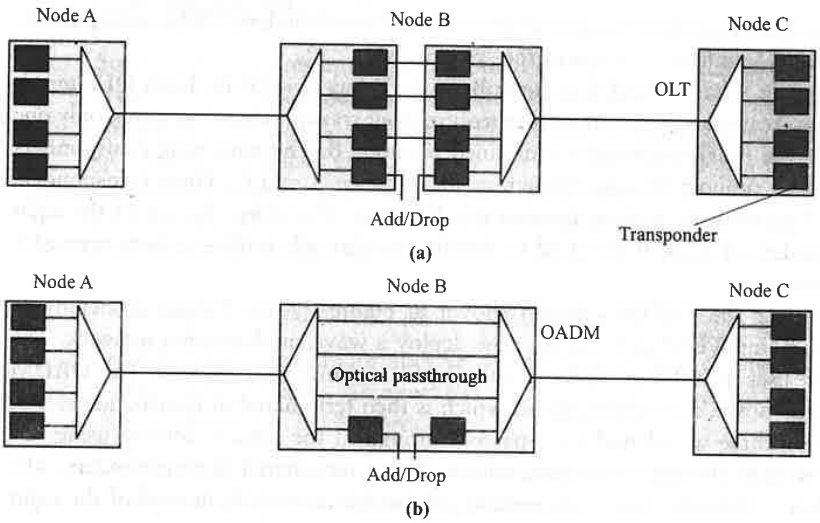

Figure 7.4 A three-node linear network example to illustrate the role of optical add/drop multiplexers. Three wavelengths are needed between nodes A and C, and one wavelength each between nodes A and B and between nodes B and C. (a) A solution using point-to-point WDM systems. (b) A solution using an optical add/drop multiplexer at node B.

amplifier sites in long-haul networks but can also be used as stand-alone network elements, particularly in metro networks. To understand the benefits of OADMs, consider a network between three nodes, say, A, B, and C, shown in Figure 7.4, with IP routers located at nodes A, B, and C. This network supports traffic between A and B, B and C, and A and C. Based on the network topology, traffic between A and C passes through node B. For simplicity, we will assume full-duplex links and full-duplex connections. This is the case for most networks today. Thus the network in Figure 7.4 actually consists of a pair of fibers carrying traffic in opposite directions.

Suppose the traffic requirement is as follows: one wavelength between A and B, one wavelength between B and C, and three wavelengths between A and C. Now suppose we deploy point-to-point WDM systems to support this traffic demand. The resulting solution is shown in Figure 7.4(a). Two point-to-point systems are deployed, one between A and B and the other between B and C. As we saw earlier in Section 7.1, each point-to-point system uses an OLT at each end of the link. The

OLT includes multiplexers, demultiplexers, and transponders. These transponders constitute a significant portion of the system cost.

Consider what is needed at node B. Node B has two OLTs. Each OLT terminates four wavelengths and therefore requires four transponders. However, only one out of those four wavelengths is destined for node B. The remaining transponders are used to support the passthrough traffic between A and C. These transponders are hooked back to back to provide this function. Therefore, six out of the eight transponders at node B are used to handle passthrough traffic—a very expensive proposition.

Consider the OADM solution shown in Figure 7.4(b). Instead of deploying point-to-point WDM systems, we now deploy a wavelength-routing network. The network uses an OLT at nodes A and C and an OADM at node B. The OADM drops one of the four wavelengths, which is then terminated in transponders. The remaining three wavelengths are passed through in the optical domain using relatively simple filtering techniques, without being terminated in transponders. The net effect is that only two transponders are needed at node B, instead of the eight transponders required for the solution shown in Figure 7.4(a). This represents a significant cost reduction. We will explore this subject of cost savings in detail in Section 8.1.

In typical carrier networks, the fraction of traffic that is to be passed through a node without requiring termination can be quite large at many of the network nodes. Thus OADMs perform a crucial function of passing through this traffic in a cost-effective manner.

Going back to our example, the reader may ask why transponders are needed in the solution of Figure 7.4(a) to handle the passthrough traffic. In other words, why can't we simply eliminate the transponders and connect the WDM multiplexers and demultiplexers between the two OLTs at node B directly, as shown in Figure 7.4(b), rather than designing a separate OADM? Indeed, this is possible, provided those OLTs are engineered to support such a capability. The physical layer engineering for networks is considerably more complex than that for point-to-point systems, as we saw in Chapter 5. For example, in a simple point-to-point system design, the power level of a signal coming into node B from node A might be so low that it cannot be passed through for another hop to node C. Also, in a network, the power of the signals added at a node must ideally be equal to the power of the signals passing through. However, there are also simpler and less expensive methods for building OADMs, as we will see in Section 7.3.1.

We will see in the next section that today's OADMs are rather inflexible. They are, for the most part, static elements and do not allow in-service selection under software control of what channels are dropped and passed through. We will see how *reconfigurable* OADMs can be built in Section 7.3.2, using tunable filters and lasers.

Ex. 1016.77

7.3.1 OADM Architectures

Several architectures have been proposed for building OADMs. These architectures typically use one or more of the multiplexers/filters that we studied in Chapter 3. Most practical OADMs use either fiber Bragg gratings, dielectric thin-film filters, or arrayed waveguide gratings. Here, we view an OADM as a black box with two line ports carrying the aggregate set of wavelengths and a number of local ports, each dropping and adding a specific wavelength. The key attributes to look for in an OADM are the following:

- What is the total number of wavelengths that can be supported?
- What is the maximum number of wavelengths that can be dropped/added at the OADM? Some architectures allow only a subset of the total number of wavelengths to be dropped/added.
- Are there constraints on whether specific wavelengths can be dropped/added? Some architectures only allow a certain set of wavelengths to be dropped/added and not any arbitrary wavelength. This capability ranges from being able to add/drop a single wavelength, to groups of wavelengths, to any arbitrary wavelength. This has a significant impact on how traffic can be routed in the network, as we will see below.
- How easy is it to add and drop additional channels? Is it necessary to take a service hit (i.e., disrupt existing channels) in order to add/drop an additional channel? This is the case with some architectures but not with others.
- Is the architecture modular, in the sense that the cost is proportional to the number of channels dropped? This is important to service providers because they prefer to "pay as they grow" as opposed to incurring a high front-end cost. In other words, service providers usually start with a small number of channels in the network and add additional channels as traffic demands increase.
- What is the complexity of the physical layer (transmission) path design with the OADM and how does adding new channels or nodes affect this design? Fundamentally, if the overall passthrough loss seen by the channels is independent of the number of channels dropped/added, then adding/dropping additional channels can be done with minimal impact to existing channels. (Other impairments like crosstalk would still have to be factored in, however.) This is an important aspect of the design that we will pay close attention to.
- Is the OADM reconfigurable, in the sense that selected channels can be dropped/added or passed through under remote software control? This is a desirable feature to minimize manual intervention. For instance, if we need to drop an

Ex. 1016.78

additional channel at a node due to traffic growth at that node, it would be simpler to do so under remote software control rather than sending a craftsperson to that location. We will study this issue in Section 7.3.2.

Figure 7.5 shows three different OADM architectures, and Table 7.1 compares their salient attributes. Several other variants are possible, and some will be explored in Problem 7.1.

In the parallel architecture (Figure 7.5(a)), all incoming channels are demultiplexed. Some of the demultiplexed channels can be dropped locally and others are passed through. An arbitrary subset of channels can be dropped and the remaining passed through. So there are no constraints on what channels can be dropped and added. As a consequence this architecture imposes minimal constraints on planning lightpaths in the network. In addition, the loss through the OADM is fixed, independent of how many channels are dropped and added. So if the other transmission impairments discussed in Chapter 5 are taken care of by proper design, then adding and dropping additional channels does not affect existing channels. Unfortunately, this architecture is not very cost-effective for handling a small number of dropped channels because, regardless of how many channels are dropped, all channels need to be demultiplexed and multiplexed back together. Therefore we need to pay for all the demultiplexing and multiplexing needed for all channels, even if we need to drop only a single channel. This also results in incurring a higher loss through the OADM. However, the architecture becomes cost-effective if a large fraction of the total number of channels is to be dropped, or if complete flexibility is desired with respect to adding and dropping any channel. The other impact of this architecture is that since all channels are demultiplexed and multiplexed at all the OADMs, each lightpath passes through many filters before reaching its destination. As a result, wavelength tolerances on the multiplexers and lasers (see Section 5.6.6) can be fairly stringent.

Some cost improvements can be made by making the design modular as shown in Figure 7.5(b). Here, the multiplexing and demultiplexing is done in two stages. The first stage of demultiplexing separates the wavelengths into bands, and the second stage separates the bands into individual channels. For example, a 16-channel system might be implemented using four bands, each having 4 channels. If only 4 channels are to be dropped at a location, the remaining 12 channels can be expressed through at the band level, instead of being demultiplexed down to the individual channel level. In addition to the cost savings in the multiplexers and demultiplexers realized, the use of bands allows signals to be passed through with lower optical loss and better loss uniformity. Several commercially available OADMs use this approach. Moreover, as the number of channels becomes large, a modular multistage multiplexing approach (see Section 3.3.10) becomes essential. Parallel OADMs are

Ex. 1016.79

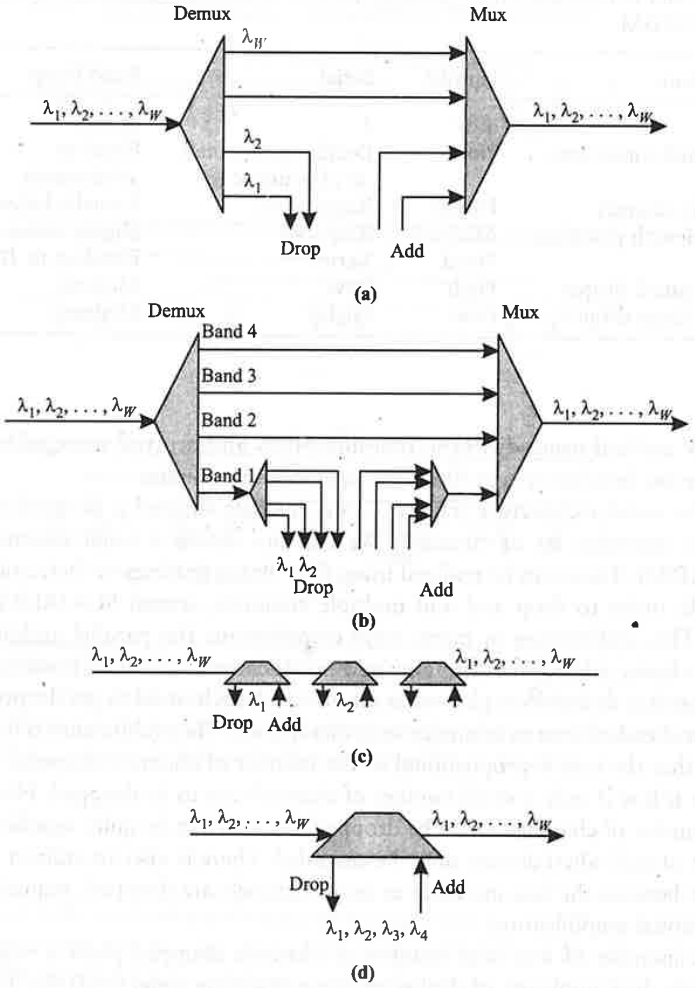

Figure 7.5 Different OADM architectures. (a) Parallel, where all the wavelengths are separated and multiplexed back; (b) modular version of the parallel architecture; (c) serial, where wavelengths are dropped and added one at a time; and (d) band drop, where a band of wavelengths are dropped and added together. W denotes the total number of wavelengths.

Table 7.1 Comparison of different OADM architectures. W is the total number of channels and D represents the maximum number of channels that can be dropped by a single OADM.

Attribute	Parallel	Serial	Band Drop
D	$= W$	1	$\ll W$
Channel constraints	None	Decide on channels at planning stage	Fixed set of channels
Traffic changes	Hitless	Requires hit	Partially hitless
Wavelength planning	Minimal	Required	Highly constrained
Loss	Fixed	Varies	Fixed up to D
Cost (small drops)	High	Low	Medium
Cost (large drops)	Low	High	Medium

typically realized using dielectric thin-film filters and arrayed waveguide gratings, and may use interleaver-type filters for large channel counts.

In the serial architecture (Figure 7.5(c)), a single channel is dropped and added from an incoming set of channels. We call this device a single-channel OADM (SC-OADM). These can be realized using fiber Bragg gratings or dielectric thin-film filters. In order to drop and add multiple channels, several SC-OADMs are cascaded. This architecture in many ways complements the parallel architecture described above. Adding and dropping additional channels disrupts existing channels. Therefore it is desirable to plan what set of wavelengths need to get dropped at each location ahead of time to minimize such disruptions. The architecture is highly modular in that the cost is proportional to the number of channels dropped. Therefore the cost is low if only a small number of channels are to be dropped. However, if a large number of channels are to be dropped, the cost can be quite significant since a number of individual devices must be cascaded. There is also an indirect impact on the cost because the loss increases as more channels are dropped, requiring the use of additional amplification.

The increase of loss with number of channels dropped plays a major role in increasing the complexity of deploying networks using serial OADMs. This is illustrated by the simple example shown in Figure 7.6. Suppose the allowed link budget for a lightpath between a transmitter and a receiver is 25 dB. Consider a situation where a lightpath from node B to node D is deployed with a loss of close to 25 dB between its transmitter and receiver. Now consider the situation when a new lightpath is to be supported at a different wavelength from node A to node C. In order to support this lightpath, an additional SC-OADM must be deployed at node C (and at node A) to drop the new lightpath. This OADM introduces an additional loss, say,

Ex. 1016.81

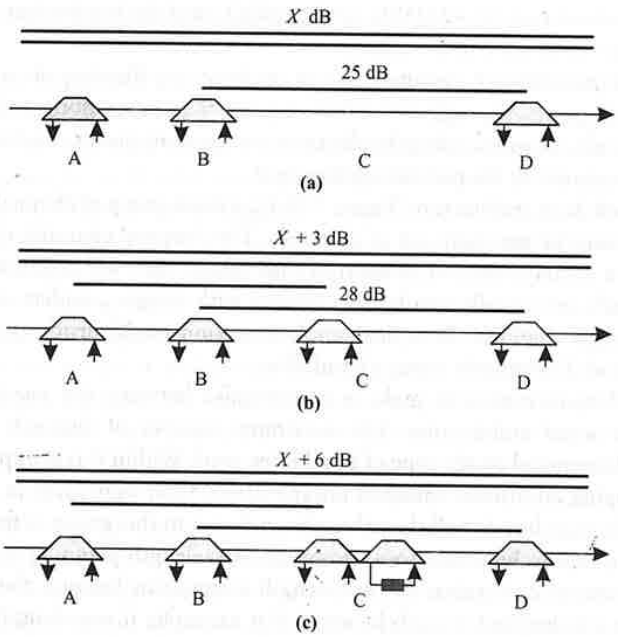

Figure 7.6 Impact of traffic changes on a network using serial OADMs. (a) Initial situation. (b) A new lightpath is added between node A and node C, causing lightpath BD to fail. (c) Lightpath BD is regenerated by adding a regenerator at node C. However, this causes other lightpaths flowing through C to be impacted.

of 3 dB, to the channels passing through node C. Introducing this OADM suddenly increases the loss on the lightpath from B to D to 28 dB, making it inoperative. The story doesn't end there, however! Suppose that in order to fix this problem we decide to regenerate this lightpath at node C. In order to regenerate this lightpath, we need to drop it at node C, send it through a regenerator, and add it back. This requires an additional SC-OADM at node C, which introduces 3 dB of additional loss for channels passing through node C. This in turn could disrupt other lightpaths passing through node C. Therefore adding or dropping additional channels can have a ripple effect on all the other lightpaths in the network. The use of optical amplifiers in conjunction with careful link engineering can alleviate some of these problems. For instance, a certain amount of loss can be allocated up front, after an optical

Ex. 1016.82

amplifier is introduced. SC-OADMs can be added until the loss budget is met, after which another amplifier can be added.

Note that passthrough channels do not undergo any filtering. As a result, each lightpath only passes through two filters, one at the source node and one at the destination node. Thus wavelength tolerances on the multiplexers and lasers are less stringent, compared to the parallel architecture.

In the band drop architecture (Figure 7.5(d)), a fixed group of channels is dropped and added from the aggregate set of channels. The dropped channels then typically go through a further level of demultiplexing where they are separated out. The added channels are usually combined together with simple couplers and added to the passthrough channels. A typical implementation could drop, say, 4 adjacent channels out of 32 channels using a band filter.

This architecture tries to make a compromise between the parallel architecture and the serial architecture. The maximum number of channels that can be dropped is determined by the type of band filter used. Within this group of channels, adding/dropping additional channels doesn't affect other lightpaths in the network as the passthrough loss for all the other channels not in this group is fixed.

However, this architecture does complicate wavelength planning in the network and places several constraints on wavelength assignment because the same set of wavelengths are dropped at each location. For example, if wavelength λ_1 is added at a node and dropped at the next node, all other wavelengths, say, λ_2, λ_3, λ_4, in the same band as λ_1 will also be added at the same node and dropped at the next node. What makes this not so ideal is that once a wavelength is dropped as part of a band, it will likely need to be regenerated before it can be added back into the network. So in this example, wavelengths λ_2, λ_3, λ_4 will need to be regenerated at both nodes even if they are passing through. It is difficult to engineer the link budget to allow optical passthrough of these wavelengths without regeneration. This problem can be fixed by using different varieties of OADMs, each of which drops a different set of channels. As the reader can readily imagine, this makes network planning complicated. If wavelength drops can be planned in advance and the network remains static, then this may be a viable option. However, in networks where the traffic changes over time, this may not be easy to plan.

The architectures that we discussed above are the ones that are feasible based on today's technology, and commercial implementations of all of these exist today. It is clear that none of them offers a perfect solution that meets a full range of applications. Serial and band drop architectures have a low entry cost, but their deployment has been hindered due to the lack of flexibility in dealing with traffic changes in the network. There is certainly a trend toward building parallel architectures, while trying to retain a reasonable initial cost.

Ex. 1016.83

7.3.2 Reconfigurable OADMs

Reconfigurability is a very desirable attribute in an OADM. Reconfigurability refers to the ability to select the desired wavelengths to be dropped and added on the fly, as opposed to having to plan ahead and deploy appropriate equipment. This allows carriers to be flexible when planning their network and allows lightpaths to be set up and taken down dynamically as needed in the network. The architectures that we considered in Figure 7.5 were not reconfigurable in this sense.

Figure 7.7 shows a few different reconfigurable OADM architectures. Figure 7.7(a) shows a variation of the parallel architecture. It uses optical switches to add/drop specific wavelengths as and when needed. Figure 7.7(b) shows a variation of the serial architecture where each SC-OADM is now a tunable device that is capable of either dropping and adding a specific wavelength, or passing it through.

Both of these architectures only partially address the reconfigurability problem because transponders are still needed to provide the adaptation into the optical layer. We distinguish between two types of transponders: a fixed-wavelength transponder and a tunable transponder. A fixed-wavelength transponder is capable of transmitting and receiving at a particular fixed wavelength. This is the case with most of the transponders today. A tunable transponder, on the other hand, can be set to transmit at any desired wavelength and receive at any desired wavelength. A tunable transponder uses a tunable WDM laser and a broadband receiver capable of receiving any wavelength.

With fixed-wavelength transponders, in order to make use of the reconfigurable OADMs shown in Figure 7.7(a) and (b), we need to deploy the transponders ahead of time so that they are available when needed. This leads to two problems: First, it is expensive to have a transponder deployed and not used while the associated OADM is passing that wavelength through. But let us suppose that this cost is offset by the added value of being able to set up and take down lightpaths rapidly. The second problem is that although the OADMs are reconfigurable, the transponders are not. So we still need to decide ahead of time as to which set of wavelengths we will need to deploy transponders for, making the network planning problem more constrained.

Avoiding these problems requires the use of tunable transponders, and even more flexible architectures than the ones shown in Figure 7.7(a) and (b). For example, Figure 7.7(c) shows a serial architecture where we have full reconfigurability. Each tunable SC-OADM is capable of adding/dropping *any* single wavelength and passing the others through, as opposed to a fixed wavelength. The adaptation is performed using a tunable transponder. This provides a fully reconfigurable OADM. Likewise, Figure 7.7(d) shows a parallel architecture with full reconfigurability. Note that this

Ex. 1016.84

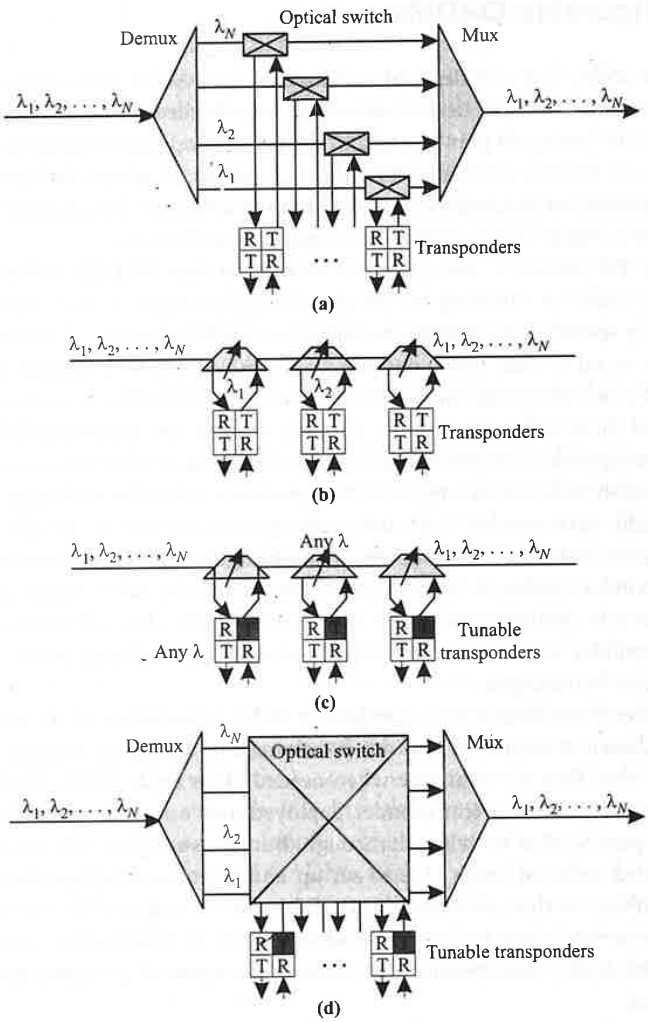

Figure 7.7 Reconfigurable OADM architectures. (a) A partially tunable OADM using a parallel architecture with optical add/drop switches and fixed-wavelength transponders. T indicates a transmitter and R indicates a receiver. (b) A partially tunable OADM using a serial architecture with fixed-wavelength transponders. (c) A fully tunable OADM using a serial architecture with tunable transponders. This transponder uses a tunable laser (marked T in the shaded box) and a broadband receiver. (d) A fully tunable OADM using a parallel architecture with tunable transponders.

Ex. 1016.85

architecture requires the use of a large optical switch. This is exactly the optical crossconnect that we will study next.

So what would an ideal OADM look like? Such an OADM (1) would be capable of being configured to drop a certain maximum number of channels, (2) would allow the user to select what specific channels are dropped/added and what are passed through under remote software control, including the transponders, without affecting the operation of existing channels, (3) would not require the user to plan ahead as to what channels may need to be dropped at a particular node, and (4) would maintain a low fixed loss regardless of how many channels are dropped/added versus passed through. The architecture of Figure 7.7(d) meets these criteria but may not be suitable for small-sized nodes where only a few channels need be dropped, due to its relatively high up-front cost. Other architectures will emerge as new component technologies such as tunable add/drops and tunable lasers become mature.

7.4 Optical Crossconnects

OADMs are useful network elements to handle simple network topologies, such as the linear topology shown in Figure 7.4 or ring topologies, and a relatively modest number of wavelengths. An additional network element is required to handle more complex mesh topologies and large numbers of wavelengths, particularly at hub locations handling a large amount of traffic. This element is the optical crossconnect (OXC). We will see that though the term "optical" is used, an OXC could internally use either a pure optical or an electrical switch fabric. An OXC is also the key network element enabling reconfigurable optical networks, where lightpaths can be set up and taken down as needed, without having to be statically provisioned.

Consider a large carrier central office hub. This might be an office in a large city for local service providers or a large node in a long-haul service provider's network. Such an office might terminate several fiber links, each carrying a large number of wavelengths. A number of these wavelengths might not need to be terminated in that location but rather passed through to another node. The OXC shown in Figure 7.8 performs this function. OXCs work alongside SONET/SDH network elements as well as IP routers and ATM switches, and WDM terminals and add/drop multiplexers as shown in Figure 7.8. Typically some OXC ports are connected to WDM equipment and other OXC ports to terminating devices such as SONET/SDH ADMs, IP routers, or ATM switches. Thus, the OXC provides cost-effective passthrough for express traffic not terminating at the hub as well as collects traffic from attached equipment into the network. Some people think of an OXC as a crossconnect switch together with the surrounding OLTs. However, our definition of OXC doesn't include

Ex. 1016.86

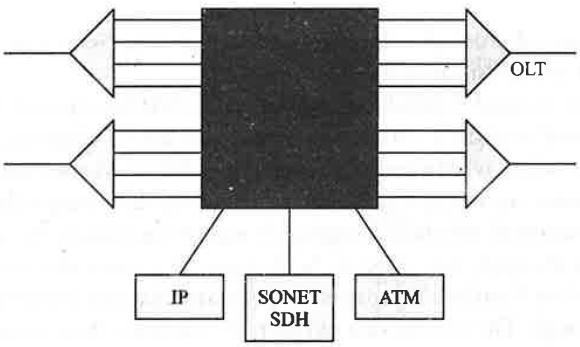

Figure 7.8 Using an OXC in the network. The OXC sits between the client equipment of the optical layer and the optical layer OLTs.

the surrounding OLTs, because carriers view crossconnects and OLTs as separate products and often buy OXCs and OLTs from different vendors.

An OXC provides several key functions in a large network:

Service provisioning. An OXC can be used to provision lightpaths in a large network in an automated manner, without having to resort to performing manual patch panel connections. This capability becomes important when we deal with large numbers of wavelengths in a node or with a large number of nodes in the network. It also becomes important when the lightpaths in the network need to be reconfigured to respond to traffic changes. The manual operation of sending a person to each office to implement a patch panel connection is expensive and error prone. Remotely configurable OXCs take care of this function.

Protection. Protecting lightpaths against fiber cuts and equipment failures in the network is emerging as one of the most important functions expected from a crossconnect. The crossconnect is an intelligent network element that can detect failures in the network and rapidly reroute lightpaths around the failure. Crossconnects enable true mesh networks to be deployed. These networks can provide particularly efficient use of network bandwidth, compared to the SONET/SDH rings we discussed in Chapter 6. We discuss this topic in detail in Chapter 10.

Bit rate transparency. The ability to switch signals with arbitrary bit rates and frame formats is a desirable attribute of OXCs.

Performance monitoring, test access, and fault localization. OXCs provide visibility to the performance parameters of a signal at intermediate nodes. They usually

allow test equipment to be hooked up to a dedicated test port where the signals passing through the OXC can be monitored in a non-intrusive manner. Non-intrusive test access requires *bridging* of the input signal. In bridging, the input signal is split into two parts. One part is sent to the core, and the other part is made available at the test access port. OXCs also provide loopback capabilities. This allows a lightpath to be looped back at intermediate nodes for diagnostic purposes.

Wavelength conversion. In addition to switching a signal from one port to another port, OXCs may also incorporate wavelength conversion capabilities.

Multiplexing and grooming. OXCs typically handle input and output signals at optical line rates. However, they can incorporate multiplexing and grooming capabilities to switch traffic internally at much finer granularities, such as STS-1 (51 Mb/s). Note that this time division multiplexing has to be done in the electrical domain and is really SONET/SDH multiplexing, but incorporated into the OXC, rather than in a separate SONET/SDH box.

An OXC can be functionally divided into a switch core and a port complex. The switch core houses the switch that performs the actual crossconnect function. The port complex houses port cards that are used as interfaces to communicate with other equipment. The port interfaces may or may not include optical-to-electrical (O/E) or optical-to-electrical-to-optical (O/E/O) converters.

Figure 7.9 shows different types of OXCs and different configurations for interconnecting OXCs with OLTs or OADMs in a node. The scenarios differ in terms of whether the actual switching is done electrically or optically, in the use of O/E and O/E/O converters, and how the OXC is interconnected to the surrrounding equipment. Table 7.2 summarizes the main differences between these architectures.

The first three configurations shown in Figure 7.9 are *opaque* configurations—the optical signal is converted into the electrical domain as it passes through the node. The last configuration (Figure 7.9(d)) is an *all-optical* configuration—the signal remains in the optical domain as it passes through the node.

Looking at Figure 7.9, observe that in the opaque configurations the switch core can be electrical or optical; that is, signals may be switched either in the electrical domain or in the optical domain. An electrical switch core can groom traffic at fine granularities and typically includes time division multiplexing of lower-speed circuits into the line rate at the input and output ports. Today, we have electrical core OXCs switching signals at granularities of STS-1 (51 Mb/s) or STS-48 (2.5 Gb/s). In contrast, a true optical switch core does not offer any grooming. It simply switches signals from one port to another.

An electrical switch core is designed to have a total switch capacity, for instance, 1.28 Tb/s. This capacity can be utilized to switch, say, up to 512 OC-48 (2.5 Gb/s)

Ex. 1016.88

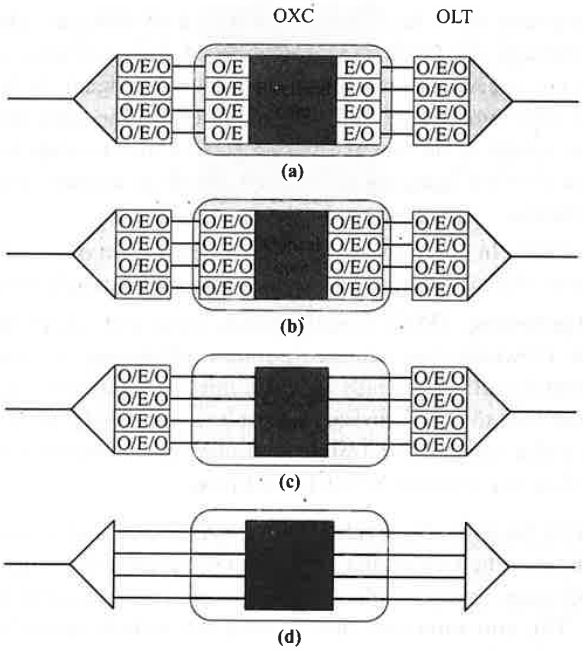

Figure 7.9 Different scenarios for OXC deployment. (a) Electrical switch core; (b) optical switch core surrounded by O/E/O converters; (c) optical switch core directly connected to transponders in WDM equipment; and (d) optical switch core directly connected to the multiplexer/demultiplexer in the OLT. Only one OLT is shown on either side in the figure, although in reality an OXC will be connected to several OLTs.

signals or 128 OC-192 (10 Gb/s) signals. The optical core is typically bit rate independent. Therefore a 1000-port optical switch core can switch 1000 OC-48 streams, 1000 OC-192 streams, or even 1000 OC-768 (40 Gb/s) streams, all at the same cost per port. The optical core is thus more scalable in capacity, compared to an electrical core, making it more future proof as bit rates increase in the future. In particular, the configuration of Figure 7.9(d) allows us to switch groups of wavelengths or all the wavelengths on a fiber together on a single OXC port, making that configuration capable of handling enormous overall capacities, and reducing the number of OXC ports required in a node.

As bit rates increase, the cost of a port on an electrical switch increases. For instance, an OC-192 port might cost twice as much as an OC-48 port. The cost of

Ex. 1016.89

Table 7.2 Comparison of different OXC configurations. Some configurations use optical to electrical converters as part of the crossconnect, in which case, they are able to measure electrical layer parameters such as the bit error rate (BER) and invoke network restoration based on this measurement. For the first two configurations, the interface on the OLTs is typically a SONET short-reach (SR), or very-short-reach (VSR) interface. For the opaque photonic configuration, it is an intermediate-reach (IR) or a special VSR interface. The cost, power, and footprint comparisons are made based on characteristics of commercially available equipment at OC-192 line rates.

Attribute	Opaque Electrical Figure 7.9(a)	Opaque Optical with O/E/Os Figure 7.9(b)	Opaque Optical Figure 7.9(c)	All-Optical Figure 7.9(d)
Low-speed grooming	Yes	No	No	No
Switch capacity	Low	High	High	Highest
Wavelength conversion	Yes	Yes	Yes	No
Switching triggers	BER	BER	Optical power	Optical power
Interface on OLT	SR/VSR	SR/VSR	IR/serial VSR	Proprietary
Cost per port	Medium	High	Medium	Low
Power consumption	High	High	Medium	Low
Footprint	High	High	Medium	Low

a port on an optical core switch, on the other hand, is the same regardless of the bit rate. Therefore at higher bit rates, it will be more cost-effective to switch signals through an optical core OXC than an electrical core OXC.

An optical switch core is also transparent; it does not care whether it is switching a 10 Gb/s Ethernet signal or a 10 Gb/s SONET signal. In contrast, electrical switch cores require separate port cards for each interface type, which convert the input signal into a format suitable for the switch fabric.

Figure 7.9(a) shows an OXC consisting of an electrical switch core surrounded by O/E converters. The OXC interoperates with OLTs through standard non-WDM short-reach (SR) optical interfaces, typically at 1310 nm. We are also seeing the deployment of cheaper very-short-reach (VSR) interfaces. The OLT has transponders to convert this signal into the appropriate WDM wavelength. Alternatively the OXC itself may have wavelength-specific lasers that operate with the OLTs without requiring transponders between them.

Figure 7.9(b)–(d) show OXCs with an optical switch core. The differences between the figures lie in how they interoperate with the WDM equipment. In Figure 7.9(b), the interworking is done in a somewhat similar fashion as in Figure 7.9(a)—through the use of O/E/O converters with short-reach or very-short-reach optical interfaces between the OXC and the OLT. In Figure 7.9(c), there are no O/E/O

Ex. 1016.90

converters and the optical switch core directly interfaces with the transponders in the OLT. Figure 7.9(d) shows a different scenario where there are no transponders in the OLT and the wavelengths in the fiber are directly switched by the optical switch core in the OXC after they are multiplexed/demultiplexed. The cost, power, and overall node footprint all improve as we go from Figure 7.9(b) to Figure 7.9(d). The electrical core option typically uses higher power and takes up more footprint, compared to the optical option, but the relative cost depends on how the different products are priced, as well as the operating bit rate on each port.

The OXCs in Figure 7.9(a) and (b) both have access to the signals in the electrical domain and can therefore perform extensive performance monitoring (signal identification and bit error rate measurements). The bit error rate measurement can also be used to trigger protection switching. Moreover, they can signal to other network elements by using inband overhead channels embedded in the data stream. (We will study signaling in more detail in Chapter 9.)

The OXCs in Figure 7.9(c) and (d) do not have the capability to look at the signal, and therefore they cannot do extensive signal performance monitoring. Thus, they cannot, for instance, invoke protection switching based on bit error rate monitoring, but instead could use optical power measurement as a trigger. These crossconnects need an out-of-band signaling channel to exchange control information with other network elements. With the configuration of Figure 7.9(c), the attached equipment needs to have optical interfaces that can deal with the loss introduced by the optical switch. These interfaces will also need to be single-mode fiber interfaces since that is what most optical switches are designed to handle. In addition, serial interfaces (single fiber pair) are preferred rather than parallel interfaces (multiple fiber pairs), as each fiber pair consumes a port on the optical switch.

The all-optical configuration of Figure 7.9(d) provides a truly all-optical network. However, it mandates a more complex physical layer design (see Chapter 5) as signals are now kept in the optical domain all the way from their source to their destination, being switched optically at intermediate nodes. Given that link engineering is complex and usually vendor proprietary, it is not easy to have one vendor's OXC interoperate with another vendor's OLT in this configuration.

Note also that the configurations of Figure 7.9(b), (c), and (d) can all be combined in a single OXC. We could have some ports having O/E/Os, others connected to OLTs with O/E/Os, and still others connected to OLTs without any O/E/Os.

It is possible to integrate the OXC and OLT systems together into one piece of equipment. Doing so provides some significant benefits. It eliminates the need for redundant O/E/Os in multiple network elements, allows tight coupling between the two to support efficient protection, and makes it easier to signal between multiple OXCs in a network using the optical supervisory channel available in the OLTs. For

Ex. 1016.91

example, in Figure 7.9(a), we could have WDM interfaces directly on the crossconnect and eliminate the intraoffice short-reach interface. We would migrate from the configuration in Figure 7.9(b) to the configuration in Figure 7.9(c).

However, this integration also has the drawback of making it a single-vendor solution. Service providers must then buy all their WDM equipment, including OLTs and OXCs, from the same vendor in order to realize this simplification. Some service providers prefer to build their network by mixing and matching "best-in-class" equipment from multiple vendors. Moreover, this solution doesn't address the problem of dealing with legacy situations where the OLTs are already deployed and OXCs must be added later.

7.4.1 All-Optical OXC Configurations

We now focus the discussion on understanding some of the issues associated with the all-optical configuration of Figure 7.9(d). As shown, the configuration can be highly cost-effective relative to the other configurations, but lacks three key functions: low-speed grooming, wavelength conversion, and signal regeneration. Low-speed grooming is needed to aggregate the lower-speed traffic streams properly for transmission over the fiber. Optical signals need to be regenerated once they have propagated through a number of fiber spans and/or other lossy elements.

Wavelength conversion is needed to improve the utilization of the network. We illustrate this with the simple example shown in Figure 7.10. Each link in the three-node network can carry three wavelengths. We have two lightpaths currently set up on each link in the network as shown and need to set up a new lightpath from node A to node C. Figure 7.10(a) shows the case where node B cannot perform wavelength conversion. Even though there are free wavelengths available in the network, the same wavelength is not available on both links in the network. As a result, we cannot set up the desired lightpath. On the other hand, if node B can convert wavelengths, then we can set up the lightpath as shown in Figure 7.10(b).

Note the configurations of Figure 7.9(a), (b), and (c) all provide wavelength conversion and signal regeneration either in the OXC itself or by making use of the transponders in the attached OLTs. Figure 7.9(a) also provides low-speed grooming, assuming that the electrical core has been designed to support that capability. In order to provide grooming, signal regeneration, and wavelength conversion, the configuration of Figure 7.9(d) is modified to include an electrical core crossconnect as shown in Figure 7.11. This configuration allows most of the signals to be switched in the optical domain, minimizing the cost and maximizing the capacity of the network, while allowing us to route the signals down to the electrical layer whenever necessary. As we discussed earlier, we could save optical switch ports by switching signals through in wavelength bands or even entire fibers at a time.

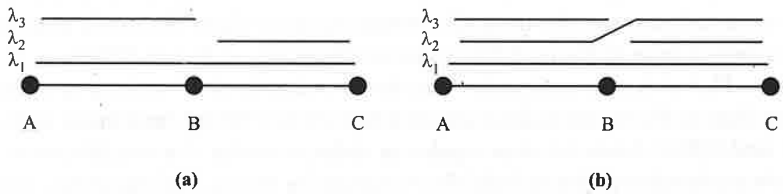

Figure 7.10 Illustrating the need for wavelength conversion. (a) Node B does not convert wavelengths. (b) Node B can convert wavelengths.

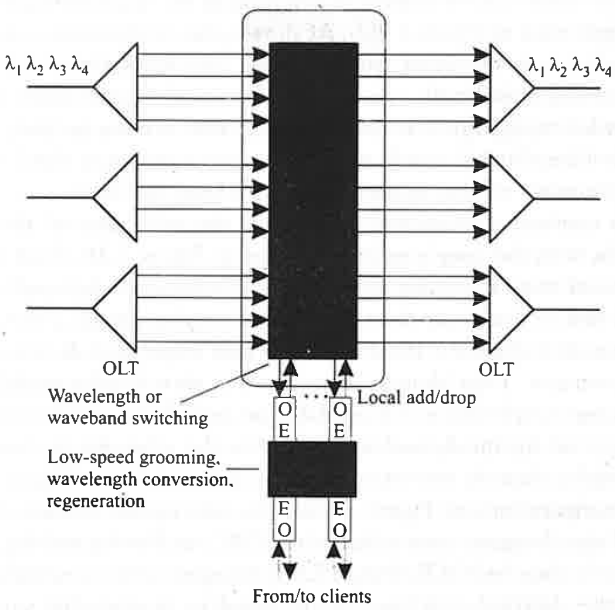

Figure 7.11 A realistic "all-optical" network node combining optical core crossconnects with electrical core crossconnects. Signals are switched in the optical domain whenever possible but routed down to the electrical domain whenever they need to be groomed, regenerated, or converted from one wavelength to another.

Ex. 1016.93

Looking at Figure 7.11, note that the optical switch does not have to switch signals from any input port to any output port. For example, it does not need to switch a signal entering at wavelength λ_1 to an output port that is connected to a multiplexer that takes in wavelength λ_2. This allows some potential simplification by making use of *wavelength planes*.

Figure 7.12 shows a wavelength plane OXC. The signals coming in over different fiber pairs are first demultiplexed by the OLTs. All the signals at a given wavelength are sent to a switch dedicated to that wavelength, and the signals from the outputs of the switches are multiplexed back together by the OLTs. In a node with F WDM fiber pairs and W wavelengths on each fiber pair, this arrangement uses F OLTs and W $2F \times 2F$ switches. This allows any or all signals on any input wavelength to be dropped locally. In contrast, the configuration of Figure 7.11 uses F OLTs and a $2WF \times 2WF$ switch to provide the same capabilities. Consider, for example, $F = 4$, $W = 32$, which are realistic numbers today. In this case, the configuration of Figure 7.12 uses four OLTs and 32 8×8 switches. In contrast, Figure 7.9(b) requires four OLTs and a 256×256 switch. As we saw in Section 3.7, larger optical switches are significantly harder to build than small ones and will need to use technologies like analog beam-steering micromirrors, whereas small optical switches can be realized using a variety of different technologies.

Based on the discussion above, it would appear that the wavelength plane approach offers a cheaper alternative to large-scale nonblocking optical switches. However, we did not consider how to optimize the number of add/drop terminations (which would be transponders or O/E interfaces on electrical switch cores). Both Figure 7.11 and Figure 7.12 assume that there are sufficient ports to terminate all WF signals. This is almost never the case, as only a fraction of traffic will need to be dropped, and the terminations are expensive. Moreover, observe that if we indeed do need WF terminations on an electrical switch, the best solution is to use the electrical core configuration of Figure 7.9(a), without having the wavelength plane switches!

If we have a total of T terminations, with all of them having tunable lasers, and we would like to drop any of the WF signals, this requires an additional $T \times WF$ optical switch between the wavelength plane switches and the terminations, as shown in Figure 7.13. In contrast, with a large nonblocking switch, we would simply connect the T terminations to T ports of this switch, resulting in a $(WF + T) \times (WF + T)$ switch overall. This situation somewhat reduces the appeal of a wavelength plane approach.

To summarize, the wavelength plane approach needs to take into account the number of fibers, fraction of add/drop traffic, number of terminations, and their tuning capabilities as separate parameters in the design. With a large-scale switch, we can partition the ports in a flexible way to account for variations in all these

Ex. 1016.94

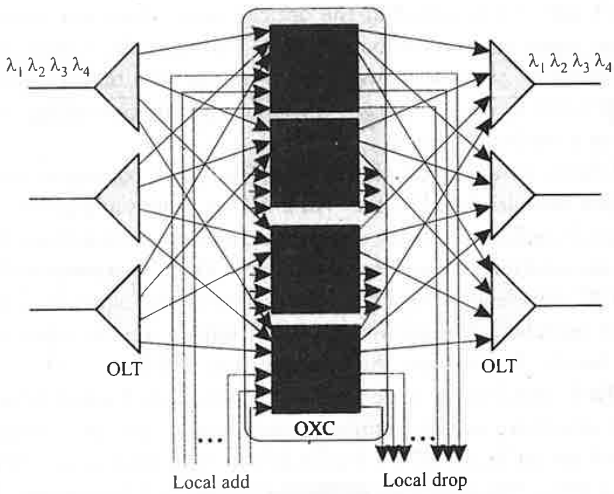

Figure 7.12 An optical core wavelength plane OXC, consisting of a plane of optical switches, one for each wavelength. With F fibers and W wavelengths on each fiber, each switch is a $2F \times 2F$ switch, if we want the flexibility to drop and add any wavelength at the node.

parameters—the only constraint is in the total number of ports available. See Problem 7.7 for another example of these types of trade-offs.

As of this writing, both electrical core and optical core OXCs are becoming available. Electrical core OXCs with total capacities up to a few Tb/s, capable of grooming down to STS-1 (51 Mb/s), are becoming available. Optical core OXCs with over 1000 ports are also emerging as commerical products, and wavelength plane OXCs are being offered by some vendors as well.

Summary

We studied the basic network elements constituting WDM networks in this chapter. We refer the reader back to Chapter 3 to get an understanding of the various technologies that are used to build these elements.

The WDM network provides circuit-switched lightpaths that can have varying degrees of transparency associated with them. Wavelengths can be reused in the network to support multiple lightpaths as long as no two lightpaths are assigned the same wavelength on a given link. Lightpaths may be protected by the network in

Ex. 1016.95

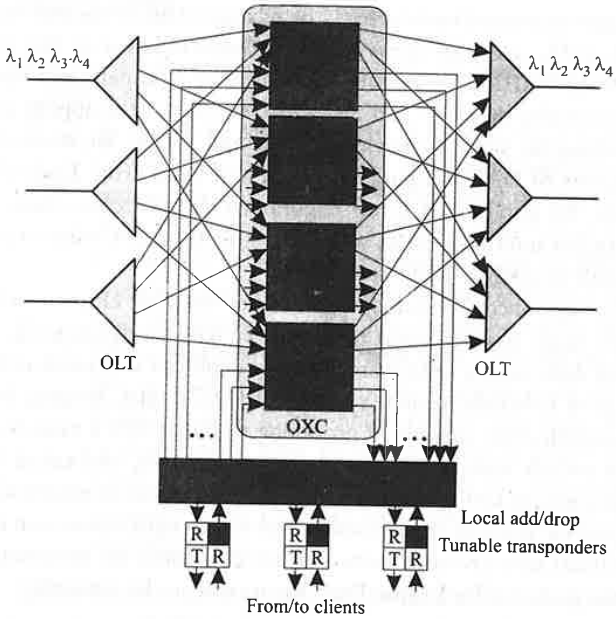

Figure 7.13 Dealing with add/drop terminations in a wavelength plane approach. An additional optical switch is required between the tunable transponders and the wavelength plane switches. Here, T denotes a transmitter, assumed to be a tunable transmitter on the WDM side, and R denotes a receiver.

the event of failures. Lightpaths can be used to provide flexible interconnections between users of the optical network, such as IP routers, allowing the router topology to be tailored to the needs of the router network.

An optical line terminal (OLT) multiplexes and demultiplexes wavelengths and is used for point-to-point applications. It typically includes transponders, multiplexers, and optical amplifiers. Transponders provide the adaptation of user signals into the optical layer. They also consitute a significant portion of the cost and footprint in an OLT. In some cases, transponders can be eliminated by deploying interfaces that provide already-adapted signals at the appropriate wavelengths in other equipment.

An optical add/drop multiplexer (OADM) drops and adds a selective number of wavelengths from a WDM signal, while allowing the remaining wavelengths to pass through. OADMs provide a cost-effective way of performing this function, compared

Ex. 1016.96

to using OLTs interconnected back to back, or relying on other equipment to handle the passthrough traffic. OADMs are typically deployed in linear or ring topologies.

Several types of OADMs are possible with a range of capabilities based on the number of wavelengths they can add and drop, the ease of dropping and adding additional wavelengths, static or reconfigurable, and so on. We studied the basic architectural flavors of OADMs: parallel, serial, and band drop. Each of these has its pros and cons. We also looked at reconfigurable OADM architectures, which use tunable filters and/or multiplexers, as well as tunable lasers, in order to provide the maximum possible flexibility in the network.

An optical crossconnect (OXC) is the other key network element in the optical layer. OXCs are large switches used to provision services dynamically as well as provide network restoration. OXCs are typically deployed in a mesh network configuration. As with OADMs, several variants of OXCs exist, ranging from OXCs with electrical switch cores capable of grooming traffic at STS-1 rates to all-optical OXCs that can switch wavelengths, bands of wavelengths, and entire fibers. Optical core crossconnects can also be surrounded by optical-to-electrical-to-optical converters to provide some of the grooming and wavelength conversion capabilities offered by electrical core crossconnects, but are not suited for grooming traffic at fine granularities such as STS-1 rates. Each has its role in the network.

Further Reading

Information regarding the various types of OLTs, OADMs, and OXCs is not easy to come by because many of the commercial implementations are proprietary in nature. Browsing through network equipment vendors' Web pages will provide some illustration of the capabilities of the different products in this space. However, there are several early testbeds that explored various forms of these network elements. For instance, [Ale93, Kam96] used a static all-optical OXC that provided a fixed interconnection pattern without any switching or wavelength conversion. [Cha94, CEG+96] explored a parallel WADM architecture as well as an OXC with an electrical switch core. [Hil93] developed an all-optical OXC without wavelength conversion. [WASG96, Gar98] developed a parallel WADM as well as a small all-optical OXC without wavelength conversion. See also [HH96, OWS96, Der95, MS96, Ber96a, Ber96b, Bac96, RS95, Chb98, KWK+98] for other relevant testbeds and architectures. The use of wavelength bands has been discussed in various contexts in [Ste90, GRW00, SS99]. For a discussion of optical crossconnects and a comparison of them to electrical crossconnects, see [GR00, GRL00].

Ex. 1016.97

Problems

7.1 Consider a ring network with two intermediate adjacent nodes A and B, each with an OADM.

(a) Consider the case where the OADM at node A adds wavelength λ_1 and the OADM at node B drops the adjacent wavelength λ_2. Suppose the minimum received power is set at −30 dBm and the transmit power is set at 0 dBm. Adjacent channel crosstalk at the receiver must be less than 15 dB. Assume that signals are added and dropped by the OADMs with no loss. What is the crosstalk suppression required at the OADM for the adjacent channel? How does this change with the link loss between the two nodes?

(b) Next consider the case where both OADMs drop and add wavelength λ_1. We are worried about the case where some of the λ_1 power, instead of being dropped at the node, "leaks" through. The intrachannel crosstalk at the receiver must be at least −30 dB below the desired signal. For the same assumptions as above, what is the intrachannel crosstalk suppression required at the OADM? How does this change with the link loss between the two nodes?

7.2 This problem illustrates some of the difficulties facing network planners when they have to use OADMs that are constrained in what channels they can add and drop. Consider a four-node linear network with nodes A, B, C, and D in that order. We have three wavelengths $\lambda_1, \lambda_2, \lambda_3$ available and are given OADMs that drop two fixed channels. That is, we can put in OADMs that drop either λ_1, λ_2, or λ_2, λ_3, or λ_1, λ_3. Now consider the situation where we need to set up the following lightpaths: AB, BC, CD, AC, BD. What OADMs would you deploy at each of the nodes? Suppose at a later point the lightpath traffic changes and now we need to replace lightpaths AC and BD by AD and BC. What changes would you have to make to support this new traffic?

7.3 Consider a linear network with serial OADMs. Assume that the transmit power is 0 dBm; the minimum received power is −30 dBm; and each OADM has a passthrough loss of 2 dB, a loss of 1 dB for the drop path, and a loss of 1 dB for the add path. Assume that the adjacent channel suppression offered by each OADM is 20 dB and that at the receiver the adjacent channel power must be at least 15 dB less than the desired signal power.

(a) Write a computer program that takes as its input the set of lightpaths in the network and their wavelengths, the loss between each pair of adjacent nodes, and determines whether each lightpath is feasible or not. The program should also determine any wavelength conflicts, that is, if two lightpaths overlap and are assigned the same wavelength.

(b) What is the maximum number of OADMs that a lightpath can pass through before it needs to be regenerated? Plot this number as a function of the total link loss in the network.

(c) Plot the output of your program for a network with five nodes numbered sequentially from 1 to 5, link loss of 5 dB between nodes, and the following sets of lightpaths and wavelength assignments: $(1, 5, \lambda_1)$, $(2, 4, \lambda_2)$, $(3, 5, \lambda_3)$, $(1, 4, \lambda_4)$.

7.4 This problem explores architectures for constructing fully reconfigurable OADMs. Consider the parallel architecture shown in Figure 7.7(d). How would you build a fully reconfigurable parallel OADM like this one, wihtout using a large optical switch? You are allowed to use tunable filters, passive splitters and combiners, and small (2×2) optical switches. These solutions need to meet properties (1) and (2) specified for the ideal OADM described in Section 7.3. With respect to property (3), you still need to keep the loss fixed, regardless of how many channels are dropped or added, but are allowed to have a reasonably high value for this loss. Compare the pros and cons of your solution versus the one in Figure 7.7(d).

7.5 You have to design a five-node ring network with a hub node and four remote nodes. Each remote node needs two wavelengths of traffic to/from the hub node on both sides of the ring; that is, you will need to dedicate two wavelengths to each remote node and terminate all the wavelengths at the hub node. You have to pick between two systems.

The first system uses eight channels in two bands, each with four channels. It provides band OADMs, which can drop one out of the two bands. Once a band is dropped, all four wavelengths in the band have to be regenerated. A band OADM costs $20,000 and a single-channel regenerator costs $10,000. No optical amplifiers are required with this system.

The second system also uses eight channels but has no bands. It provides SC-OADMs, which can drop any single wavelength. Each SC-OADM costs $10,000. For this system, two optical line amplifiers are required, each costing $30,000. Whose system would you select based on just equipment cost?

7.6 This problem illustrates the need for large OXCs and also illustrates the value of using wavelength bands.

Consider an all-optical OXC with 256 ports deployed in the configuration shown in Figure 7.9(d). Each WDM line system carries 32 wavelengths; 75% of the lightpaths pass through the node while the remaining 25% are dropped and added onto routers attached to the OXC. Each lightpath added and dropped onto a router takes up two OXC ports.

(a) How many WDM line systems can the OXC support?

Ex. 1016.99

(b) Next suppose that 25% of the lightpaths passing through need to be converted from one wavelength to another. This is done by sending the lightpath to one of a pool of regenerators/wavelength converters attached to the OXC. Each such regenerator uses two ports in the OXC. Thus a lightpath needing to be converted uses four OXC ports. Now, how many WDM line systems can the OXC support?

(c) Now suppose the WDM line systems are designed with eight bands, each with four wavelengths. Assume that all the lightpaths passing through can be passed through at the band level without having to be demultiplexed down to the individual channel. For lightpaths that are dropped and added, the entire band is dropped and demultiplexed after the bands are passed through the OXC. No wavelength conversion is needed. How many WDM line systems can the OXC support?

7.7 Consider the wavelength plane switch architecture of Figure 7.12. Consider the situation where we have a total of four fibers and 40 wavelengths on each fiber. We must design the node such that any four signals can be dropped. (Note that this implies we could potentially drop all the wavelengths on a particular fiber while passing through all the wavelengths on the other fibers.) The wavelengths are dropped onto transponders, which have tunable lasers. The 40 wavelengths are split into five bands of 8 wavelengths each, and a tunable laser can tune over a single band.

(a) Draw a block diagram of this node and indicate the minimum number of transponders needed. Compare this against an approach using large non-blocking switches.

(b) Now suppose we have tunable lasers that can tune over two bands instead of one. How does the situation change?

References

[Ale93] S. B. Alexander et al. A precompetitive consortium on wide-band all-optical networks. *IEEE/OSA Journal on Lightwave Technology*, 11:714–735, May/June 1993.

[Bac96] E.-J. Bachus et al. Coherent optical systems implemented for business traffic routing and access: The RACE COBRA project. *IEEE/OSA JLT/JSAC Special Issue on Multiwavelength Optical Technology and Networks*, 14(6):1309–1319, June 1996.

[Ber96a] L. Berthelon et al. Experimental assessment of node cascadability in a reconfigurable survivable WDM ring network. In *Proceedings of Topical Meeting on Broadband Optical Networks*, 1996.

[Ber96b] L. Berthelon et al. Over 40,000 km across a layered network by recirculation through an experimental WDM ring network. In *Proceedings of European Conference on Optical Communication*, 1996.

[CEG+96] G. K. Chang, G. Ellinas, J. K. Gamelin, M. Z. Iqbal, and C. A. Brackett. Multiwavelength reconfigurable WDM/ATM/SONET network testbed. *IEEE/OSA JLT/JSAC Special Issue on Multiwavelength Optical Technology and Networks*, 14(6):1320–1340, June 1996.

[Cha94] G. K. Chang et al. Experimental demonstration of a reconfigurable WDM/ATM/SONET multiwavelength network testbed. In *OFC'94 Technical Digest*, 1994. Postdeadline paper PD9.

[Chb98] M. W. Chbat et al. Towards wide-scale all-optical networking: The ACTS optical pan-European network (OPEN) project. *IEEE JSAC: Special Issue on High-Capacity Optical Transport Networks*, 16(7):1226–1244, Sept. 1998.

[Der95] F. Derr. Design of an 8 × 8 optical cross-connect switch: Results on subsystems and first measurements. In *ECOC'95 Optical Networking Workshop*, 1995. Paper S2.2.

[Gar98] L. D. Garrett et al. The MONET New Jersey demonstration network. *IEEE JSAC: Special Issue on High-Capacity Optical Transport Networks*, 16(7):1199–1219, Sept. 1998.

[GR00] J. Gruber and R. Ramaswami. Moving towards all-optical networks. *Lightwave*, 34(8):40–49, Dec. 2000.

[GRL00] J. Gruber, P. Roorda, and F. Lalonde. The photonic switch crossconnect (PSX)—its role in evolving optical networks. In *Proceedings of National Fiber Optic Engineers Conference*, pages 678–689, 2000.

[GRW00] O. Gerstel, R. Ramaswami, and W-K. Wang. Making use of a two stage multiplexing scheme in a WDM network. In *OFC 2000 Technical Digest*, pages ThD1-1–ThD1-3, 2000.

[HH96] A. M. Hill and A. J. N. Houghton. Optical networking in the European ACTS programme. In *OFC'96 Technical Digest*, pages 238–239, San Jose, CA, Feb. 1996.

[Hil93] G. R. Hill et al. A transport network layer based on optical network elements. *IEEE/OSA Journal on Lightwave Technology*, 11:667–679, May/June 1993.

[Kam96] I. P. Kaminow et al. A wideband all-optical WDM network. *IEEE JSAC/JLT Special Issue on Optical Networks*, 14(5):780–799, June 1996.

[KWK+98] M. Koga, A. Watanabe, T. Kawai, K. Sato, and Y. Ohmori. Large-capacity optical path cross-connect system for WDM photonic transport network. *IEEE JSAC: Special Issue on High-Capacity Optical Transport Networks*, 16(7):1260–1269, Sept. 1998.

Ex. 1016.101

[MS96] W. C. Marra and J. Schesser. Africa ONE: The Africa optical network. *IEEE Communications Magazine*, 34(2):50–57, Feb. 1996.

[OWS96] S. Okamoto, A. Watanabe, and K.-I. Sato. Optical path cross-connect node architectures for photonic transport network. *IEEE/OSA JLT/JSAC Special Issue on Multiwavelength Optical Technology and Networks*, 14(6):1410–1422, June 1996.

[RS95] R. Ramaswami and K. N. Sivarajan. Routing and wavelength assignment in all-optical networks. *IEEE/ACM Transactions on Networking*, pages 489–500, Oct. 1995. An earlier version appeared in *Proceedings of IEEE Infocom'94*.

[SS99] A. A. M. Saleh and J. M. Simmons. Architectural principles for optical regional and metropolitan access networks. *IEEE/OSA Journal on Lightwave Technology*, 17(12), Dec. 1999.

[Ste90] T. E. Stern. Linear lightwave networks: How far can they go? In *Proceedings of IEEE Globecom*, pages 1866–1872, 1990.

[WASG96] R. E. Wagner, R. C. Alferness, A. A. M. Saleh, and M. S. Goodman. MONET: Multiwavelength optical networking. *IEEE/OSA JLT/JSAC Special Issue on Multiwavelength Optical Technology and Networks*, 14(6):1349–1355, June 1996.