ULTRAFAST OPTICS

ANDREW M. WEINER Purdue University



A JOHN WILEY & SONS, INC., PUBLICATION

ULTRAFAST OPTICS

WILEY SERIES IN PURE AND APPLIED OPTICS

Founded by Stanley S. Ballard, University of Florida

EDITOR: Glenn Boreman, University of Central Florida, CREOL & FPCE

BARRETT AND MYERS · Foundations of Image Science BEISER · Holographic Scanning BERGER-SCHUNN · Practical Color Measurement BOYD · Radiometry and The Detection of Optical Radiation BUCK · Fundamentals of Optical Fibers, Second Edition CATHEY · Optical Information Processing and Holography CHUANG • Physics of Optoelectronic Devices DELONE AND KRAINOV · Fundamentals of Nonlinear Optics of Atomic Gases DERENIAK AND BOREMAN · Infrared Detectors and Systems DERENIAK AND CROWE · Optical Radiation Detectors DE VANY · Master Optical Techniques ERSOY · Diffraction, Fourier Optics and Imaging GASKILL · Linear Systems, Fourier Transform, and Optics GOODMAN · Statistical Optics HOBBS · Building Electro-Optical Systems: Making It All Work, Second Edition HUDSON · Infrared System Engineering IIZUKA · Elements of Photonics, Volume I: In Free Space and Special Media IIZUKA · Elements of Photonics, Volume II: For Fiber and Integrated Optics JUDD AND WYSZECKI · Color in Business, Science, and Industry, Third Edition KAFRI AND GLATT · The Physics of Moire Metrology KAROW · Fabrication Methods for Precision Optics KLEIN AND FURTAK · Optics, Second Edition MALACARA · Optical Shop Testing, Third Edition MILONNI AND EBERLY · Lasers NASSAU · The Physics and Chemistry of Color: The Fifteen Causes of Color, Second Edition NIETO-VESPERINAS · Scattering and Diffraction in Physical Optics OSCHE · Optical Detection Theory for Laser Applications O'SHEA · Elements of Modern Optical Design **OZAKTAS** • The Fractional Fourier Transform PRATHER, SHI, SHARKAWY, MURAKOWSKI, AND SCHNEIDER • Photonic Crystals: Theory, Applications, and Fabrication SALEH AND TEICH · Fundamentals of Photonics, Second Edition SCHUBERT AND WILHELMI · Nonlinear Optics and Quantum Electronics SHEN · The Principles of Nonlinear Optics UDD · Fiber Optic Sensors: An Introduction for Engineers and Scientists UDD · Fiber Optic Smart Structures VANDERLUGT · Optical Signal Processing VEST · Holographic Interferometry VINCENT · Fundamentals of Infrared Detector Operation and Testing WEINER · Ultrafast Optics WILLIAMS AND BECKLUND · Introduction to the Optical Transfer Function WYSZECKI AND STILES · Color Science: Concepts and Methods, Quantitative Data and Formulae, Second Edition XU AND STROUD · Acousto-Optic Devices YAMAMOTO · Coherence, Amplification, and Quantum Effects in Semiconductor Lasers YARIV AND YEH · Optical Waves in Crystals YEH · Optical Waves in Layered Media YEH · Introduction to Photorefractive Nonlinear Optics YEH AND GU · Optics of Liquid Crystal Displays

ULTRAFAST OPTICS

ANDREW M. WEINER Purdue University



A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2009 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permission.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Weiner, Andrew Marc

Ultrafast optics / Andrew M. Weiner.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-471-41539-8 (cloth)

1. Laser pulses, Ultrashort. 2. Laser pulses, Ultrashort–Industrial applications.
3. Mode-locked lasers. 4. Nonlinear optical spectroscopy. 5. Nonlinear optics. I. Title.
QC689.5.L37W45 2009

621.36'6-dc22

2008052027

Printed in the United States of America

 $10\ 9\ 8\ 7\ 6\ 5\ 4\ 3\ 2\ 1$

In memory of

Jason Weiner, my father

Prof. Hermann Haus

and

Bill Drake, Jr.

CONTENTS

1 Introduction and Review

	1.1 Introduction to Ultrashort Laser Pulses, 1		
	1.2 Brief Review of Electromagnetics, 4		
	1.2.1 Maxwell's Equations, 4		
	1.2.2 The Wave Equation and Plane Waves, 6		
	1.2.3 Poynting's Vector and Power Flow, 8		
	1.3 Review of Laser Essentials, 10		
	1.3.1 Steady-State Laser Operation, 10		
	1.3.2 Gain and Gain Saturation in Four-Level Atoms, 15		
	1.3.3 Gaussian Beams and Transverse Laser Modes, 17		
	1.4 Introduction to Ultrashort Pulse Generation Through Mode-Locking, 22		
	1.5 Fourier Series and Fourier Transforms, 25		
	1.5.1 Analytical Aspects, 25		
	1.5.2 Computational Aspects, 28		
	Problems, 30		
2	Principles of Mode-Locking		
	2.1 Processes Involved in Mode-Locking, 32		
	2.2 Active Mode-Locking, 33		
	2.2.1 Time-Domain Treatment, 34		
	2.2.2 Frequency-Domain Treatment, 40		
	2.2.3 Variations of Active Mode-Locking, 43		
	2.3 Passive Mode-Locking Using Saturable Absorbers, 44		
	2.3.1 Saturation Model, 47		

- 2.3.2 Slow Saturable Absorber Mode-Locking, 50
- 2.3.3 Fast Saturable Absorber Mode-Locking, 54

xiii

- viii CONTENTS
 - 2.4 Solid-State Laser Mode-Locking Using the Optical Kerr Effect, 57
 - 2.4.1 Nonlinear Refractive Index Changes, 57
 - 2.4.2 Self-Amplitude Modulation, Self-Phase Modulation, and Group Velocity Dispersion, 58
 - 2.4.3 Additive Pulse Mode-Locking, 60
 - 2.4.4 Kerr Lens Mode-Locking, 64
 - 2.4.5 Mode-Locking Solutions, 75
 - 2.4.6 Initiation of Mode-Locking, 81
 - Problems, 83

3 Ultrafast-pulse Measurement Methods

- 3.1 Terminology and Definitions, 85
- 3.2 Electric Field Autocorrelation Measurements and the Power Spectrum, 88
- 3.3 Electric Field Cross-Correlation Measurements and Spectral Interferometry, 91
 - 3.3.1 Electric Field Cross-Correlation, 92
 - 3.3.2 Spectral Interferometry, 93
 - 3.3.3 Application: Optical Coherence Tomography, 96
- 3.4 Intensity Correlation Measurements, 99
 - 3.4.1 Correlation Measurements Using Second-Harmonic Generation, 99
 - 3.4.2 Experimental Procedures, 108
 - 3.4.3 Correlation Measurements Using Two-Photon absorption, 110
 - 3.4.4 Higher-Order Correlation Techniques, 111
- 3.5 Chirped Pulses and Measurements in the Time–Frequency Domain, 112
- 3.6 Frequency-Resolved Optical Gating, 118
 - 3.6.1 Polarization-Gating FROG, 119
 - 3.6.2 Self-Diffraction FROG, 122
 - 3.6.3 Second-Harmonic-Generation FROG, 124
 - 3.6.4 Frequency-Resolved Optical Gating Using Temporal Phase Modulation, 125
 - 3.6.5 Signal Recovery from FROG Traces, 126
- 3.7 Pulse Measurements Based on Frequency Filtering, 130
 - 3.7.1 Single-Slit Approaches, 131
 - 3.7.2 Double-Slit Approach, 134
- 3.8 Self-Referencing Interferometry, 135
 - 3.8.1 Time-Domain Interferometry of Chirped Pulses, 135
 - 3.8.2 Self-Referencing Spectral Interferometry, 137
- 3.9 Characterization of Noise and Jitter, 139
- Problems, 144

4 Dispersion and Dispersion Compensation

- 4.1 Group Velocity Dispersion, 147
 - 4.1.1 Group Velocity Definition and General Dispersion Relations, 147
 - 4.1.2 General Aspects of Material Dispersion, 151
- 4.2 Temporal Dispersion Based on Angular Dispersion, 155
 - 4.2.1 Relation Between Angular and Temporal Dispersion, 155
 - 4.2.2 Angular Dispersion and Tilted Intensity Fronts, 159
- 4.3 Dispersion of Grating Pairs, 161

147

- 4.4 Dispersion of Prism Pairs, 166
- 4.5 Dispersive Properties of Lenses, 173
- 4.6 Dispersion of Mirror Structures, 177
 - 4.6.1 The Gires–Tournois Interferometer, 178
 - 4.6.2 Quarter-Wave Stack High Reflectors, 180
 - 4.6.3 Chirped Mirrors, 182
- 4.7 Measurements of Group Velocity Dispersion, 1864.7.1 Interferometric Methods, 187
 - 4.7.1 Interferometric Methods, 187
 - 4.7.2 Frequency-Domain Intracavity Dispersion Measurements, 190
- 4.8 Appendix, 191
 - 4.8.1 Frequency-Dependent Phase Due to Propagation Through a Slab: Alternative Derivation, 191

4.8.2 Impedance Method for Analysis of Dielectric Mirror Stacks, 192 Problems, 195

5 Ultrafast Nonlinear Optics: Second Order

- 5.1 Introduction to Nonlinear Optics, 198
- 5.2 The Forced Wave Equation, 201
 - 5.2.1 Frequency-Domain Formulation, 202
 - 5.2.2 Time-Domain Formulation, 203
- 5.3 Summary of Continuous-Wave Second-Harmonic Generation, 204
 - 5.3.1 Effect of Phase Matching, 207
 - 5.3.2 Phase Matching in Birefringent Media, 209
 - 5.3.3 Focusing Effects in Continuous-Wave SHG, 215
- 5.4 Second-Harmonic Generation with Pulses, 220
 - 5.4.1 SHG in the Quasi-Continuous-Wave Limit, 220
 - 5.4.2 Ultrashort-Pulse SHG, 221
 - 5.4.3 Quasi-Phase Matching, 228
 - 5.4.4 Effect of Group Velocity Walk-off on SHG-Based Pulse Measurements, 233
- 5.5 Three-Wave Interactions, 237
 - 5.5.1 Sum Frequency Generation, 240
 - 5.5.2 Difference Frequency Generation, 244
 - 5.5.3 Optical Parametric Amplification, 245
- 5.6 Appendix, 253
 - 5.6.1 Spatial Walk-off and Pulse Fronts in Anisotropic Media, 253
 - 5.6.2 Velocity Matching in Broadband Noncollinear Three-Wave Mixing, 254
- Problems, 256

6 Ultrafast Nonlinear Optics: Third Order

- 6.1 Propagation Equation for Nonlinear Refractive Index Media, 258
 - 6.1.1 Plane Waves in Uniform Media, 260
 - 6.1.2 Nonlinear Propagation in Waveguides, 261
 - 6.1.3 Optical Fiber Types, 264
- 6.2 The Nonlinear Schrödinger Equation, 266
- 6.3 Self-Phase Modulation, 270
 - 6.3.1 Dispersionless Self-Phase Modulation, 270
 - 6.3.2 Dispersionless Self-Phase Modulation with Loss, 273

258

X CONTENTS

- 6.3.3 Self-Phase Modulation with Normal Dispersion, 274
- 6.3.4 Cross-Phase Modulation, 275
- 6.4 Pulse Compression, 276
- 6.5 Modulational Instability, 283
- 6.6 Solitons, 286
- 6.7 Higher-Order Propagation Effects, 291
 - 6.7.1 Nonlinear Envelope Equation in Uniform Media, 292
 - 6.7.2 Nonlinear Envelope Equation in Waveguides, 295
 - 6.7.3 Delayed Nonlinear Response and the Raman Effect, 296
 - 6.7.4 Self-Steepening, 306
 - 6.7.5 Space–Time Focusing, 308
- 6.8 Continuum Generation, 310

Problems, 313

7 Mode-Locking: Selected Advanced Topics

- 7.1 Soliton Fiber Lasers: Artificial Fast Saturable Absorbers, 316
 - 7.1.1 The Figure-Eight Laser, 317
 - 7.1.2 Energy Quantization, 322
 - 7.1.3 Soliton Sidebands, 324
- 7.2 Soliton Mode-Locking: Active Modulation and Slow Saturable Absorbers, 328
 - 7.2.1 Harmonically Mode-Locked Soliton Fiber Lasers, 328
 - 7.2.2 The Net Gain Window in Soliton Mode-Locking, 330
- 7.3 Stretched Pulse Mode-Locking, 337
 - 7.3.1 Stretched Pulse Mode-Locked Fiber Laser, 337
 - 7.3.2 Dispersion-Managed Solitons, 340
 - 7.3.3 Theoretical Issues, 342
- 7.4 Mode-Locked Lasers in the Few-Cycle Regime, 344
- 7.5 Mode-Locked Frequency Combs, 347
 - 7.5.1 Comb Basics, 347
 - 7.5.2 Measurement Techniques, 350
 - 7.5.3 Stabilization of Frequency Combs, 354
 - 7.5.4 Applications, 356
- Problems, 360

8 Manipulation of Ultrashort Pulses

- 8.1 Fourier Transform Pulse Shaping, 362
 - 8.1.1 Examples of Pulse Shaping Using Fixed Masks, 364
 - 8.1.2 Programmable Pulse Shaping, 369
 - 8.1.3 Pulse-Shaping Theory, 376
- 8.2 Other Pulse-Shaping Techniques, 386
 - 8.2.1 Direct Space-to-Time Pulse Shaping, 386
 - 8.2.2 Acousto-optic Dispersive Filters, 390
- 8.3 Chirp Processing and Time Lenses, 394
 - 8.3.1 Space–Time Duality, 394
 - 8.3.2 Chirp Processing, 397
 - 8.3.3 Time Lens Processing, 399

- 8.4 Ultrashort-Pulse Amplification, 405
 - 8.4.1 Amplification Basics, 406
 - 8.4.2 Special Issues in Femtosecond Amplifiers, 411
- 8.5 Appendix, 416
 - 8.5.1 Fresnel Diffraction and Fourier Transform Property of a Lens, 416
 - 8.5.2 Wave Optics Model of a Grating, 418
- Problems, 420

9 Ultrafast Time-Resolved Spectroscopy

- 9.1 Introduction to Ultrafast Spectroscopy, 422
- 9.2 Degenerate Pump–Probe Transmission Measurements, 426
 - 9.2.1 Co-polarized Fields: Scalar Treatment, 426
 - 9.2.2 Vector Fields and Orientational Effects, 431
- 9.3 Nondegenerate and Spectrally Resolved Pump–Probe: Case Studies, 439
 - 9.3.1 Femtosecond Pump–Probe Studies of Dye Molecules, 440
 - 9.3.2 Femtosecond Pump–Probe Studies of GaAs, 444
- 9.4 Basic Quantum Mechanics for Coherent Short-Pulse Spectroscopies, 451
 - 9.4.1 Some Basic Quantum Mechanics, 451
 - 9.4.2 The Density Matrix, 456
- 9.5 Wave Packets, 460
 - 9.5.1 Example: Semiconductor Quantum Wells, 461
 - 9.5.2 Molecules, 462
- 9.6 Dephasing Phenomena, 469
 - 9.6.1 Linear Spectroscopies, 469
 - 9.6.2 Models of Dephasing, 475
 - 9.6.3 Measurement of Dephasing Using Transient Gratings, 481
 - 9.6.4 Two-Dimensional Spectroscopy, 494
- 9.7 Impulsive Stimulated Raman Scattering, 499 Problems, 505

10 Terahertz Time-Domain Electromagnetics

- 10.1 Ultrafast Electromagnetics: Transmission Lines, 507
 - 10.1.1 Photoconductive Generation and Sampling, 507
 - 10.1.2 Electro-optic Sampling, 513
- 10.2 Ultrafast Electromagnetics: Terahertz Beams, 516
 - 10.2.1 Generation and Measurement of Terahertz Pulses, 517
 - 10.2.2 Terahertz Spectroscopy and Imaging, 527

Problems, 531

References

533

507

PREFACE

This book deals with the optics of picosecond and femtosecond light pulses, primarily at wavelengths in the visible range and longer. Research in ultrafast optics started roughly forty years ago, and although this field is now tremendously active, in many aspects it has also reached a level of maturity. However, relatively few broad treatments of ultrafast optics are available. It is hoped that this book, which is both detailed and comprehensive, will be a valuable resource not only for graduate students and researchers seeking to enter ultrafast optics but also for colleagues already engaged in this fascinating field.

I would like to mention a few pertinent points about my perspective in this book. First, in keeping with my training as an electrical engineer, the signals aspect of ultrafast optics is emphasized. That is, I often attempt to capture the detailed form of ultrashort pulses as they are transformed in various optical systems or evolve inside mode-locked lasers. Similarly, the detailed form of measurement data, as in ultrashort pulse characterization and ultrafast spectroscopy, is analyzed when possible.

Second, although a detailed theoretical treatment is often presented, I strive to balance this with an experimental perspective. Accordingly, many examples of data from the literature are included, especially in the later chapters. These examples are selected to provide concrete illustration of material that otherwise might remain abstract, to provide evidence of certain important phenomena, and sometimes to illustrate applications.

Third, although the suite of applications of ultrafast optics is now very rich, this book is concerned primarily with fundamental principles. No attempt is made to cover the applications space comprehensively. Two applications are covered in some depth: ultrafast spectroscopy and ultrafast electromagnetic pulse generation and measurement. Both of these are subjects of individual chapters. Certain other applications, such as the application of optical frequency combs for precision frequency metrology, are discussed briefly within appropriate sections of related text.

Fourth, the book is focused on ultrafast optics in visible and lower-frequency spectral bands, on time scales down to femtoseconds, and at intensities for which perturbative non-linear optics applies. Under these conditions the motions of bound electrons that mediate

xiv PREFACE

important laser–matter interactions may usually be viewed as instantaneous. Extreme nonlinear optics phenomena, arising from high-intensity laser–matter interactions for which the laser field reaches or exceeds the interatomic field, are not covered. One important example of extreme nonlinear optics is high harmonic generation, in which visible wavelength femtosecond pulses result in emission of photons in the vacuum ultraviolet (XUV) and soft x-ray bands. The use of high harmonic generation to realize attosecond pulses has become an active research topic within the last few years. Attosecond time scales and XUV and x-ray frequencies bring in entirely new physics that are beyond the scope of this book. Attosecond technology and science are in a stage of rapid evolution and will undoubtedly be the subject of future treatises.

The structure of the book is as follows:

- Chapter 1 begins with a brief overview and motivation, discussing key attributes of ultrashort laser pulses and some application examples. Important background material, including simple electromagnetics and laser essentials, is reviewed. The chapter continues with a phenomenological introduction to short pulse generation via mode-locking and concludes with a review of Fourier transforms, a mathematical tool essential to much of our treatment.
- Chapter 2 covers basic principles of laser mode-locking in some depth. The intent is not only to cover one of the most interesting topics at the outset, but to use the discussion on mode-locking as a physical context in which to introduce a variety of important ultrafast optical effects (e.g., dispersion, filtering, self-phase modulation), many of which are themselves treated in detail in subsequent chapters.
- Measurement of pulses on the femtosecond time scale is an important issue, since the speed required is considerably faster than that of existing photodetectors and oscilloscopes. In Chapter 3 we discuss methods for characterizing ultrashort pulses. Included are historical techniques dating back to the early years of ultrafast optics (these offer only partial information but remain in widespread use) as well as more powerful techniques offering full waveform information. The field of ultrashort pulse characterization has continued to grow, and new techniques continue to be introduced. I have not attempted to cover all the interesting measurement techniques that have been invented. My hope is that the discussion accompanying those methods that are included will prepare the reader to quickly grasp additional methods that may interest him or her.
- Dispersion is often a key limiting effect in ultrafast systems. Accordingly, in Chapter 4 we focus on dispersion and its compensation. After defining key concepts, the discussion covers material dispersion, then temporal dispersion arising from angular dispersion (including important grating and prism pair setups), and finally, dispersion effects in mirror structures. The effect of dispersion in the focusing of light by lenses is also discussed, as are methods for measurement of dispersion.
- Chapters 5 and 6 deal with ultrafast nonlinear optics. Chapter 5 emphasizes second-order nonlinear effects. After an introduction to nonlinear optics and a review of continuous-wave second-harmonic generation (SHG), new effects arising in ultra-short pulse SHG, sum and difference frequency generation, and optical parametric generation are discussed. Such effects are of primary interest in frequency conversion and pulse measurement applications. Chapter 6 focuses on refractive index (third-order) nonlinearities, which have seen very wide applications in ultrafast optics. Topics

include self-phase modulation, pulse compression, solitons, continuum generation, and propagation equations, including propagation equations relevant for pulses down to a few optical cycles.

- Chapter 7 takes advantage of material developed in earlier chapters to continue the discussion of mode-locking at a more advanced level. Included are soliton nonlinear optics phenomena observed in the mode-locking of fiber lasers, stretched pulse lasers operating in the normal dispersion regime, and soliton mode-locking of solid-state lasers with slow saturable absorbers. Important aspects of sub-10-fs laser design and stabilized frequency combs important for precision frequency metrology are also discussed.
- Chapter 8 focuses on the manipulation of ultrashort pulses. The chapter begins with detailed coverage of ultrafast Fourier optics methods that enable ultrashort pulse shaping and arbitrary waveform generation. It then treats various chirped pulse approaches for waveform manipulation and measurement, including interesting time lens approaches. Finally, femtosecond pulse amplification techniques, leading to realization of pulses with unparalleled peak power, are discussed.
- Ultrafast time-resolved spectroscopy is possibly the most widely practiced application of ultrashort light pulses. This field is highly interdisciplinary, and the number of different physical systems probed using ultrafast techniques is large. In Chapter 9 I present and analyze selected important concepts in femtosecond time-resolved spectroscopy. No attempt is made to cover all the applications or all the experimental variations. A pedagogical challenge is that students studying ultrafast optics may have very different degrees of preparation in quantum mechanics, which is needed for a detailed microscopic understanding of many of the systems studied via ultrafast techniques. Therefore, this chapter takes a two-pronged approach. First, techniques such as pumpprobe, which principally probe incoherent (phase insensitive) population relaxation processes, are treated classically, although quantum concepts such as energy levels do appear phenomenologically. Experiments on population relaxation in organic dye molecules and in direct-gap semiconductors such as GaAs are discussed for illustration. Second, after a necessarily brief (and possibly inadequate) introduction to relevant quantum mechanics, spectroscopies sensitive to coherent (phase sensitive) phenomena are discussed. These subjects are treated with the help of quantum mechanics. Topics include wave packet phenomena in semiconductors and molecules, coherent polarization effects, and measurement of dephasing. A final topic, impulsive stimulated Raman scattering, is treated in a largely classical framework.
- The final chapter deals with another important application of ultrafast optics: the generation and measurement of picosecond and subpicosecond electrical and electromagnetic transients. Both electrical signals propagating on-chip on transmission-line structures and terahertz (THz) electromagnetic radiation freely propagating in space are considered. Finally, THz time-domain spectroscopy, a technique that provides exciting capability for materials characterization and sensing in a spectral region that is difficult to access by either direct electronic or optical means, is discussed.

Several problems are provided at the end of each chapter, ranging from simple theoretical questions to practical exercises requiring numerical computation. As an example of the latter, in the Chapter 2 problems the student is asked to simulate pulse evolution through many round trips in a mode-locked laser cavity, finally arriving at the self-consistent pulse

solution. I regularly assign such numerical problems in my own course on ultrafast optics at Purdue University. Although in my experience such problems require substantial effort on the part of the student, they result in a much better understanding of the phenomena involved, not to mention improved skill in applying numerical tools such as the fast Fourier transform. For homework on pulse measurement I have frequently synthesized FROG or other data on a computer; I then distribute the data file to the class with the assignment to process the data to extract the pulse shape. (I have not included such problems in the current book, as I deemed it more expedient to let instructors generate their own pulses and corresponding data files.) The numerical problems included in the book may be used as is or may simply serve to inspire instructors to invent their own numerically oriented problems.

Authoring this book has been a project of nearly ten years. I began formal writing while on sabbatical during the 1999–2000 academic year at the Max Born Institute (MBI) for Nonlinear Optics and Ultrashort Pulse Spectroscopy in Berlin, Germany. Work continued for many years, but in a fragmented way, at my home institution. I made substantial progress toward completion during a second sabbatical during the 2006–2007 academic year in Boulder, Colorado, where I split my time at the National Institute of Standards of Technology (NIST) and at JILA, a joint enterprise of NIST and the University of Colorado. I owe great thanks to my sabbatical hosts, Prof. Thomas Elsaesser of the MBI, Dr. Leo Hollberg of NIST, and Prof. Steve Cundiff of JILA, for making these stays possible. I would also like to thank the MBI and the Alexander von Humboldt Foundation for assistance with funding during my stay in Berlin and NIST and JILA for assistance with funding during my stay in Boulder.

I would like to thank many persons who generously provided input and assistance in various aspects of this project. Giullo Cerullo, Steve Cundiff, Alex Gaeta, and Franz Kaertner provided helpful comments and clarification on various technical topics (in some cases, on multiple topics). Virginia Lorenz made available a copy of her University of Colorado Ph.D. thesis, which provided a very helpful overview of dephasing. At Purdue University, Dee Dee Dexter provided invaluable logistical and secretarial assistance throughout the course of this project. Dan Leaird was always willing to lend an ear when I wanted to voice ideas about the book project; Dan also deserves great thanks for his unflagging attention to our ultrafast optics and fiber communications research laboratory, even when I sometimes became distracted by the burdens of authorship. Many students deserve credit for identifying errors in preliminary versions of the manuscript, which were used over several iterations of my graduate course. Prof. Dongsun Seo, a sabbatical visitor from Korea, also pointed out several items in need of correction. A number of graduate students kindly agreed to carry out numerical work, generating data that resulted in a number of figures. These students include Jung-Ho Chung, Ehsan Hamidi, Zhi Jiang, Houxun Miao, Bhaskaran Muralidharan, Ninad Pimparkar, Haifeng Wang, Mark Webster, and Shang-Da Yang. V. R. Supradeepa checked several equations on my behalf. Zhi Jiang was especially helpful in proofing the typeset manuscript.

This book includes well over 200 figures, many of which were composed especially for this project. Although many others were taken from the literature, almost all of these were modified or redrawn to ensure readability and to achieve consistency of appearance and notation. I am tremendously grateful to Bill Drake, Jr., for fulfilling this responsibility with great skill from the inception of this project until it neared completion. Tragically, Bill succumbed to cancer at an early age. He continued to contribute to this book even as he struggled against the disease that ultimately killed him. Michael Black took over technical illustration responsibilities during the final period of this project; he also deserves much gratitude.

Finally, I would like to thank my parents, Jason and Geraldine Weiner, for fostering in me a love of learning; my graduate advisors, Profs. Hermann Haus and Erich Ippen, for attracting me to the field of ultrafast optics; and my wife, Brenda, and children, Roberta, Steven, and Gabriela, for their love and patience.

ANDREW M. WEINER

West Lafayette, Indiana September 2008

1

INTRODUCTION AND REVIEW

1.1 INTRODUCTION TO ULTRASHORT LASER PULSES

This book is about ultrafast laser pulses: what they are, linear and nonlinear optical effects which they experience, methods by which they are generated and measured, and how they can be used for measurement of ultrafast physical processes. Let us begin with a definition of the relevant time units.

1 nanosecond (ns) = 10^{-9} s = 0.000000001s 1 picosecond (ps) = 10^{-12} s = 0.000000000001s 1 femtosecond (fs) = 10^{-15} s = 0.000000000000001s 1 attosecond = 10^{-18} s = 0.0000000000000000001s

To put these very short time units in perspective, it is useful to consider their spatial equivalent. If we could take a snapshot of a 1-s light pulse, this pulse would stretch over a distance of 186,000 miles (or 300,000 km), equal to the speed of light multiplied by 1s. This is roughly three-fourths of the distance from the Earth to the moon, a distance we will consider very slow! Now skipping over milliseconds and microseconds, we arrive at nanoseconds. One nanosecond has a spatial extent of 30 cm (ca. 1 ft). Although still rather slow by the standards of ultrafast optics, the nanosecond is the approximate time scale for high-speed electronic chips and computers. The word *ultrafast* is usually applied to the picosecond time scale and below. A picosecond has an extent of 0.3 mm, roughly the thickness of a business card. Given that typical garden-variety laser beams have beam diameters on the order of a few millimeters, we should perhaps envision pulses a picosecond and shorter not as pencils of light but as pancakes of light! In the visible and near-infrared spectral regions, pulses as

Ultrafast Optics, By Andrew M. Weiner

Copyright © 2009 John Wiley & Sons, Inc.

short as a few femtoseconds can now be generated. The spatial extent of even a 10-fs laser pulse is only $3 \mu m$, much less than the diameter of a human hair.

Pulse durations of a few femtoseconds in the visible are approaching the fundamental pulse-width limitation of roughly one optical cycle (roughly one wavelength in spatial extent). Research into attosecond pulse generation is also under way [1]. One key theme in attosecond pulse generation is the use of highly nonlinear optical frequency-conversion methods to produce radiation at much higher frequencies (much shorter wavelengths), corresponding to extreme ultraviolet (XUV) and x-ray spectral regions. At such frequencies the duration of a single optical cycle (and hence the attainable pulse-width limit) is reduced, making attosecond pulses possible.

In this book we specifically focus on ultrafast optics in visible and lower-frequency spectral bands and on time scales down to femtoseconds. Within this time scale the motions of bound electrons that mediate important laser-matter interactions may usually be viewed as instantaneous. Conversely, attosecond time scales and XUV and x-ray frequencies bring in entirely new physics in which laser-matter interactions are sensitive to the noninstantaneous dynamics of bound electron motions. Attosecond technology and science are in a stage of rapid evolution and will undoubtedly be the subject of future books.

Ultrashort pulses have several related characteristics which make them useful for applications. These include the following:

- *High time resolution*. By definition, the pulse duration is in the picosecond or femtosecond range (or below). This provides very high time resolution for excitation and measurement of ultrafast physical processes in solid-state, chemical, and biological materials.
- *High spatial resolution*. The spatial extent of a short light pulse is given by the pulse duration multiplied by the speed of light. As noted above, for very short pulse durations, the spatial pulse length can be on the order of micrometers. This makes ultrashort pulses useful for some microscopy and imaging applications.
- *High bandwidth*. By the uncertainly principle, the product of the pulse-width times the optical bandwidth must be of order unity (or larger). As the pulse duration decreases, the bandwidth increases correspondingly. Pulses of 100 fs have bandwidths on the order of 10 terahertz (THz), and the shortest visible laser pulses contain so much of the visible spectrum that they appear white. This high-bandwidth feature can be important for optical communications as well as other applications.
- Potential for high intensity. For a given pulse energy, the peak power and peak intensity are inversely proportional to the pulse duration. Because the size (hence cost) of high-power lasers usually scales with pulse energy, femtosecond pulse technology can be used to obtain ultrahigh peak intensities at moderate energy levels. Amplified femtosecond pulses have produced peak powers up to the petawatt level (1 petawatt = 10^{15} W) and peak intensities exceeding 10^{20} W/cm².

The field of ultrafast optics has traditionally been a highly interdisciplinary one, with a wide range of applications areas. To give a flavor for the nature of application areas, we comment below on a few of the research applications.

Ultrafast Spectroscopy Time-resolved spectroscopy is a very successful and probably the most widespread application of picosecond and femtosecond laser technology. The idea

is that ultrashort laser pulses can be used to make "stop-action" measurements of ultrafast physical processes, just as high-speed (microsecond) electronic flashes have been used starting several decades ago to make such stop-action photographs of bullets traveling through apples and milk droplets splashing into milk bowls [2]. On the femtosecond time scale, macroscopic objects such as bullets and milk droplets are motionless, and therefore ultrafast spectroscopy is best applied to study microscopic processes. Examples include investigations of femtosecond interactions of photoexcited electrons and holes with each other and with lattice vibrations in semiconductor crystals, ultrafast laser-induced melting, photodisassociation and ultrafast solution dynamics of chemical species, and ultrafast internal rearrangements of the large organic molecule bacteriorhodopsin as photons absorbed in the retina initiate the first biochemical steps in the process of vision. The principles of ultrafast spectroscopy are covered in Chapter 9 with examples.

Laser-Controlled Chemistry In a research area closely related to ultrafast spectroscopy, researchers are using specially engineered femtosecond laser waveforms to try to influence the course of photoinduced chemical reactions. In addition to observing ultrafast chemical motions as in time-resolved spectroscopy, the added idea here is to control the motions that take place. Since the intrinsic time scale for nuclear motions in chemical systems is tens to hundreds of femtoseconds, femtosecond laser pulses are a natural tool in pursuing the challenging goal of laser-controlled chemistry.

Frequency Metrology Ultrashort pulses are usually emitted from lasers in the form of periodic trains, which under certain conditions can exhibit very high timing stability and long-term coherence. The spectrum of such a periodic train is a comb of up to hundreds of thousands of discrete spectral lines, which may be stabilized to permit precision measurements of optical frequencies with sub-hertz uncertainties across the optical spectrum. Such stabilized frequency combs are now widely adopted for high-precision frequency metrology and for investigations of precision optical clocks. Related topics are discussed in Section 7.5.

High-Speed Electrical Testing Testing is a key issue in the development of high-speed electronic devices and circuits. Electronic test instrumentation based on established technology is usually slower than advanced high-speed research devices. However, since even the very fastest electronic devices only reach into the picosecond range, ultrafast laser technology offers speed to spare. Thus femtosecond optical pulses have been applied to generate subpicosecond electrical pulses and to measure operation of the highest-speed electronic devices. Ultrafast electronic and measurement are discussed in Chapter 10.

Laser–Plasma Interactions Lasers with intensities of 10^{13} W/cm² and above (easily achieved using amplified femtosecond pulses) directed onto solid targets are sufficient to strip electrons from their nuclei, resulting in a laser-generated plasma. On the 100-fs time scale, the resulting free electrons do not have enough time to separate from the ionized nuclei. This provides the opportunity to study solid-density plasmas at temperatures as high as 1 million degrees.

Short-Wavelength Generation High-intensity ultrashort pulses at visible wavelengths can also be used to generate coherent short-wavelength radiation in the vacuum ultraviolet and

x-ray ranges through highly nonlinear harmonic generation processes or by pumping x-ray lasers. Coherent short-wavelength radiation may be important, for example, for imaging microscopic structures such as DNA.

Optical Communications The low-loss transmission window of optical fibers has a bandwidth comparable to that of a 100-fs pulse, and therefore ultrashort-pulse technology may play an important role in optical communications. Subpicosecond pulses have already been used for laboratory experiments demonstrating fiber optic transmission of data at Tbit/s (10^{12} bit/s) rates. Here ultrafast optics technology is important not only for pulse generation but also for signal processing, for data detection, and for the advanced metrology necessary for characterizing and optimizing ultrashort-pulse transmission [3,4]. Ultrashort pulses may also prove important in wavelength-division-multiplexing (WDM) systems in which the fiber bandwidth is carved up into different wavelength bands or channels. For WDM applications it is the large bandwidth of the ultrashort pulse (not the short duration) which is useful, since a single pulse contains enough bandwidth to produce a number of wavelength channels.

Biomedical Applications Ultrashort pulses are finding substantial applications in biomedical imaging. Attractive features include the ability to perform optical imaging within scattering media (e.g., most tissues) and to obtain high-resolution depth information. An example of such an application is discussed in Section 3.3.3. In confocal microscopy significantly improved spatial resolution has been demonstrated by relying on two-photon excitation. The ability of ultrashort pulses to provide high intensity without high pulse energy is important in the use of this technique with sensitive biological samples. In laser-assisted surgical procedures ultrashort pulses may in some cases reduce collateral tissue damage by reducing heat deposition.

Materials Processing High-power lasers are used for a variety of industrial applications, such as cutting and drilling. With continuous-wave or "long"-pulse (nanoseconds) lasers, the minimum feature size and the quality of the cut are limited by thermal diffusion of heat to areas neighboring the laser focus. With femtosecond lasers, materials processing is possible using lower pulse energies, due to the very high peak powers, which lead to new physical mechanisms. This reduces the heat deposited into the sample during the laser machining process and leads to a much cleaner cutting or drilling operation.

1.2 BRIEF REVIEW OF ELECTROMAGNETICS

Since ultrashort laser pulses are made up of light, and light is a form of electromagnetic radiation, we very briefly review Maxwell's equations, which describe all forms of electromagnetic radiation, including light. We use MKS (SI) units here and throughout the book. It is assumed that the reader is already familiar with vector calculus. For a more detailed treatment of electromagnetics, the reader is directed to textbooks on this subject [5,6].

1.2.1 Maxwell's Equations

Maxwell's equations are a set of relationships between the electric field E and magnetic field H (boldface symbols denote vectors). Inside a medium we must also consider the

Symbol	Name	Units or Numerical Value
E	Electric field	V m ⁻¹
D	Electric flux density	$\mathrm{C}~\mathrm{m}^{-2}$
Н	Magnetic field	$A m^{-1}$
В	Magnetic flux density	T (or V·s m^{-2})
Р	Polarization density	$\mathrm{C}~\mathrm{m}^{-2}$
Μ	Magnetization density	$A m^{-1}$
J	Current density	$\mathrm{A}~\mathrm{m}^{-2}$
ρ	Charge density	$C m^{-3}$
ε_0	Permittivity of free space	$8.85 \times 10^{-12} \text{ F m}^{-1} \text{ (or C V}^{-1} \text{ m}^{-1})$
μ_0	Permeability of free space	$4\pi \times 10^{-7} \text{ H m}^{-1} \text{ (or V s}^2 \text{ m}^{-1} \text{ C}^{-1})$

 Table 1.1
 Names and Units of Symbols in Maxwell's Equations

charge density ρ , current density **J**, polarization density **P**, and magnetization density **M**, and in order to include the effect of the fields on the matter, the electric and magnetic flux densities, **D** and **B**, are also introduced. Units for these quantities are given in Table 1.1. Maxwell's equations are then written as follows:

$$\nabla \cdot \mathbf{D} = \rho \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{1.2}$$

$$\nabla \times \mathbf{E} = \frac{-\partial \mathbf{B}}{\partial t} \tag{1.3}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{1.4}$$

The relations defining **D** and **B** are

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \tag{1.5}$$

$$\mathbf{B} = \mu_0 \left(\mathbf{H} + \mathbf{M} \right) \tag{1.6}$$

The constants ε_0 and μ_0 are known as the *permittivity* and *permeability* of free space, with the numerical values and units given in Table 1.1. Note also that the symbol ρ refers to the free charge density (i.e., any bound charge density associated with the polarization is not included). Similarly, the current density **J** does not include any currents associated with the motion of bound charges (changes in polarization). In free space we would have $\rho = \mathbf{J} = \mathbf{P} = \mathbf{M} = 0$.

For now we specialize to the case of a linear, isotropic, and source-free medium. By *source-free* we mean that the charge and current densities are zero ($\rho = 0$ and $\mathbf{J} = 0$). By *linear* we mean that the medium response (i.e., the polarization and magnetization) is linear in the applied fields. For the case of the electric field, we write

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E} \tag{1.7}$$

where χ_e is known as the electric susceptibility (dimensionless). Inserting into eq. (1.5), one obtains

$$\mathbf{D} = \varepsilon_0 (1 + \chi_e) \mathbf{E} = \varepsilon \mathbf{E}$$
(1.8)

The proportionality constant ε is termed the dielectric constant, with

$$\varepsilon = (1 + \chi_e) \varepsilon_0 \tag{1.9}$$

Other common terms include the relative dielectric constant ($\varepsilon/\varepsilon_0$) and the index of refraction *n*, which is commonly used in optics, where

$$n^2 = \frac{\varepsilon}{\varepsilon_0} \tag{1.10}$$

For the case of the magnetic field, we write

$$M = \chi_m H \tag{1.11}$$

where χ_m is the magnetic polarizability. Using eq. (1.6), we obtain

$$\mathbf{B} = \mu_0 \left(1 + \chi_m \right) \mathbf{H} = \mu \mathbf{H} \tag{1.12}$$

In most cases in ultrafast optics, one is interested in nonmagnetic materials, for which $\mathbf{M} = 0$. In this case of zero magnetization, one has

$$\mathbf{B} = \mu_0 \mathbf{H} \tag{1.13}$$

Equations (1.7) and (1.11) are examples of constitutive laws, which specify the response of the material to the fields. The form of these equations as written arises because we have assumed both linear and isotropic media (for nonisotropic media, one would need to replace the assumed scalar susceptibilities with tensors). We note that there are many situations in ultrafast optics where these assumptions are not valid. For example, nonlinear optical effects, which we discuss in later chapters, require by definition that \mathbf{P} be a nonlinear function of \mathbf{E} .

1.2.2 The Wave Equation and Plane Waves

We now consider electromagnetic wave propagation in linear, isotropic, source-free media. To derive the wave equation, we take the curl of eq. (1.3) and insert eq. (1.4), which, using the stated assumptions and a well-known vector identity,¹ gives the following:

$$\nabla \times \nabla \times \mathbf{E} = \nabla \left(\nabla \cdot \mathbf{E} \right) - \nabla^2 \mathbf{E} = -\mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}$$
(1.14)

¹ The identity is $\nabla \times \nabla \times \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$. Note that in Cartesian coordinates ∇^2 has a very simple form, namely $\nabla^2 \mathbf{A} = (\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)\mathbf{A}$.

Since $\nabla \cdot \mathbf{E} = 0$ also under our conditions, we obtain the wave equation

$$\nabla^2 \mathbf{E} = \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{1.15}$$

One situation of special interest is the case where the field varies in only one direction, which without loss of generality we take as the z direction. Then the wave equation becomes

$$\frac{\partial^2 \mathbf{E}}{\partial z^2} = \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{1.16}$$

The general solution takes the form

$$\mathbf{E}(z,t) = \mathbf{E}_0\left(t - \frac{z}{v}\right) \tag{1.17}$$

where \mathbf{E}_0 is a vector in the x-y plane [eq. (1.1) precludes *E* from having a *z*-component] and $v = 1/\sqrt{\mu\varepsilon}$. The solution can be verified by plugging back into the wave equation. Equation (1.17) is called a *plane-wave solution*, since the field does not vary in the transverse (*x*-*y*) plane. It also represents a traveling wave, since the field propagates in the *z* direction without changing its form. In the case of a pulsed field, $\mathbf{E}_0(t)$ represents the pulse shape. The propagation velocity is given by *v*. Note that

$$\frac{1}{\sqrt{\mu_0 \varepsilon_0}} = c \cong 2.998 \times 10^8 \,\mathrm{m \, s^{-1}} \tag{1.18}$$

is the velocity of light in free space. Therefore, for the case most common in optics where $\mu = \mu_0$, the velocity of propagation within a medium is given by

$$v = \frac{c}{n} \tag{1.19}$$

where *n* is the refractive index according to eq. (1.10). Note also that in deriving eqs. (1.14) to (1.17), we have assumed implicitly that the refractive index *n* is independent of frequency. When *n* does have a frequency dependence, this can change the propagation velocity or cause the pulse to distort during propagation. These effects are discussed in Chapter 4.

The case of a sinusoidal solution to the wave equation will be of special importance. Then eq. (1.17) takes the form

$$\mathbf{E}(z,t) = \mathbf{E}_0 \cos\left(\omega t - kz + \phi\right) \tag{1.20}$$

where \mathbf{E}_0 is now a constant vector, $\boldsymbol{\omega}$ is the angular frequency, and the propagation constant k must satisfy the dispersion relation

$$k = \omega \sqrt{\mu \varepsilon} \tag{1.21}$$

or again, assuming that $\mu = \mu_0$,

$$k = \frac{\omega n}{c} \tag{1.22}$$

The wave has a temporal oscillation period equal to $2\pi/\omega$ and a spatial period or wavelength *in the medium* given by $\lambda = 2\pi/k$. The wavelength in free space is denoted λ_0 and is given by

$$\lambda_0 = \frac{2\pi c}{\omega} \tag{1.23}$$

Equation (1.20) represents the ideal case of single frequency or monochromatic laser radiation. It can also be written in the equivalent form

$$\mathbf{E}(z,t) = \operatorname{Re}\left\{\tilde{\mathbf{E}}_{\mathbf{0}}e^{j(\omega t - kz)}\right\}$$
(1.24)

where $\text{Re}\{\dots\}$ denotes the real part and the phase ϕ has been incorporated into the complex vector $\tilde{\mathbf{E}}_0$. We refer to this form as complex notation. As we will see shortly, ultrashort light pulses are conveniently described as superpositions of sinusoidal solutions of the form (1.20) or (1.24) with different frequencies.

Finally, we note that similar solutions can be written for propagation in directions other than along z, as follows:

$$\mathbf{E}(\mathbf{r},t) = \operatorname{Re}\left\{\tilde{\mathbf{E}}_{\mathbf{0}}e^{j\left(\omega t - \mathbf{k} \cdot \mathbf{r}\right)}\right\}$$
(1.25)

Here **k** is the propagation vector; it points along the direction of propagation and its magnitude $k = |\mathbf{k}|$ still satisfies the dispersion relation (1.21).

1.2.3 Poynting's Vector and Power Flow

We also review the expressions for energy flow with electromagnetic waves. To arrive at the required formulas, we form the dot product of eq. (1.3) with **H** and subtract from this the dot product of eq. (1.4) with **E**. Using another vector identity,² we find that

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{E} \cdot \mathbf{J} = 0$$
(1.26)

We also make use of the divergence theorem,

$$\int \nabla \cdot \mathbf{A} \, dV = \int \mathbf{A} \cdot \hat{\mathbf{n}} \, dS \tag{1.27}$$

which states that the surface integral of a vector **A** over a closed surface is equal to the volume integral of $\nabla \cdot \mathbf{A}$ over the volume bounded by that surface. $\hat{\mathbf{n}}$ is the unit vector

² The identity is $\nabla \cdot \mathbf{A} \times \mathbf{B} = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$.

normal to the surface and pointing outward. The result is

$$\int \left(\mathbf{E} \times \mathbf{H}\right) \cdot \hat{\mathbf{n}} \, dS + \int dV \left\{ \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{E} \cdot \mathbf{J} \right\} = 0 \tag{1.28}$$

Finally, assuming a linear medium and substituting for **D** and **B** using eqs. (1.8) and (1.12), we obtain

$$\int \left(\mathbf{E} \times \mathbf{H}\right) \cdot \hat{\mathbf{n}} \, dS + \int dV \left\{ \frac{\partial \left(\frac{1}{2}\varepsilon \, |\mathbf{E}|^2\right)}{\partial t} + \frac{\partial \left(\frac{1}{2}\mu \, |\mathbf{H}|^2\right)}{\partial t} + \mathbf{E} \cdot \mathbf{J} \right\} = 0 \qquad (1.29)$$

Equations (1.28) and (1.29) are representations of Poynting's theorem, which describes conservation of energy in electromagnetic systems. We can identify specific meanings for each of the terms. Look at eq. (1.29), for example:

- ∫ (E × H) · n̂ dS is the net rate of energy flow out of the closed surface. It has units of power (watts). E × H is called the *Poynting vector* and has units of intensity (W/m²). It gives the power density carried by an electromagnetic wave and the direction in which power is carried.
- $\frac{1}{2}\varepsilon |\mathbf{E}|^2$ and $\frac{1}{2}\mu |\mathbf{H}|^2$ are the local energy densities (J/m³) associated with the electric and magnetic fields, respectively. $\partial/\partial t \int \frac{1}{2}\varepsilon |\mathbf{E}|^2 dV$ and $\partial/\partial t \int \frac{1}{2}\mu |\mathbf{H}|^2 dV$ represent the time rate of change of electric and magnetic field energy stored within the volume, respectively.
- ∫ E · J dV represents power dissipation or generation within the volume (in watts). When E · J is positive, this term represents power dissipation due, for example, to ohmic losses. Energy is transferred out of the fields and into the medium, typically as heat. When E · J is negative, this term represents power supplied by the currents and fed into the electromagnetic fields.

Overall, Poynting's theorem is a power balance equation, showing how changes in stored energy are accounted for by power dissipation and energy flow.

It is worth specializing once more to the case of single-frequency sinusoidal fields, with E given by eq. (1.25). The H field is obtained using eq. (1.3), with the result

$$\mathbf{H} = \operatorname{Re}\left\{\sqrt{\frac{\varepsilon}{\mu}} \frac{\mathbf{k} \times \tilde{\mathbf{E}}_{\mathbf{0}}}{k} e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}\right\}$$
(1.30)

Thus **H** is perpendicular to both **E** and **k**, and its magnitude is equal to $\sqrt{\varepsilon/\mu} |\mathbf{E}|$. The factor $\sqrt{\mu/\varepsilon}$ is termed the *characteristic impedance* of the medium, and $\sqrt{\varepsilon/\mu}$ is therefore the *admittance*.

In optics one is usually interested in the time-average power flow. This is calculated in complex notation as follows. First consider scalar functions f(t) and g(t), where

$$f(t) = \operatorname{Re}\left\{\tilde{f}e^{j\omega t}\right\}$$
 and $g(t) = \operatorname{Re}\left\{\tilde{g}e^{j\omega t}\right\}$ (1.31)

The time average of f(t)g(t) is given by

$$\langle fg \rangle = \frac{1}{2} \operatorname{Re} \left\{ \tilde{f} \, \tilde{g}^* \right\} \tag{1.32}$$

Here $\langle \cdots \rangle$ denotes the time average and * indicates a complex conjugate. Similarly, if $\mathbf{f}(\mathbf{t}), \mathbf{g}(\mathbf{t}), \tilde{\mathbf{f}}$, and $\tilde{\mathbf{g}}$ now denote vectors, the time average of $\mathbf{f} \times \mathbf{g}$ is given by

$$\langle \mathbf{f} \times \mathbf{g} \rangle = \frac{1}{2} \operatorname{Re} \left\{ \tilde{\mathbf{f}} \times \tilde{\mathbf{g}}^* \right\}$$
 (1.33)

Using these relations, the time-average Poynting vector for the plane waves of eqs. (1.25) and (1.30) becomes

$$\langle \mathbf{E} \times \mathbf{H} \rangle = \frac{1}{2} \sqrt{\frac{\varepsilon}{\mu}} \left| \tilde{\mathbf{E}}_{\mathbf{0}} \right|^2 \frac{\mathbf{k}}{k}$$
 (1.34)

where we have assumed that ε and μ are real. Power is carried along the direction of **k**. In the case of a nonmagnetic material, we can write the magnitude of the time-average Poynting vector, commonly called the *intensity I*, in the following useful form:

$$I = |\langle \mathbf{E} \times \mathbf{H} \rangle| = \frac{1}{2} \varepsilon_0 cn \left| \tilde{\mathbf{E}}_0 \right|^2$$
(1.35)

1.3 REVIEW OF LASER ESSENTIALS

We will shortly discuss in some detail methods by which lasers can be made to produce ultrashort light pulses. First, however, we give a brief and simple review of lasers in general. More detail can be found in texts on lasers, such as [7,8].

1.3.1 Steady-State Laser Operation

Schematic drawings of two simple laser geometries are shown in Fig. 1.1. Both lasers consist of a set of mirrors and a gain medium. The gain medium is an optical amplifier which coherently amplifies light passing through it. The mirrors may be curved or planar and together make up the laser cavity or resonator. The cavity is aligned so that light reflects back and forth again and again, passing along the same path every time. If we imagine even a very weak light intensity in the cavity (due to spontaneous emission from the gain medium), then for sufficiently high gain, the intensity increases from one round trip through the laser to the next, eventually resulting in an intense beam. In steady state the gain per round trip must equal the loss. Part of the light passes through the partially transmissive output coupler, and this forms the output laser beam, which can be used for experiments.



Figure 1.1 Schematic of (a) a linear cavity laser and (b) a ring laser. HR, high reflector; OC, output coupler.

In the linear or Fabry–Perot laser cavity shown in Fig. 1.1a, light passes along the same path, traveling from left to right and right to left. The light makes two passes through the gain medium per round trip. In the ring cavity shown in Fig. 1.1b, light passes through the gain medium only once per round trip. In the diagram we have assumed unidirectional operation (i.e., light is traveling around the ring in only one direction). In real ring lasers either unidirectional or bidirectional operation is possible. Both linear and ring geometries have been used in femtosecond laser design.

For our purposes we usually consider the gain medium as a black box. The physics of laser gain media is usually analyzed using quantum mechanics in courses on laser fundamentals. Here we usually stick to a classical description, although we will use the indispensable energy-level concept, which does come from quantum mechanics. Common laser media used for ultrafast lasers include impurity-doped crystals or glasses such as Ti:sapphire, Nd: YAG or Nd: glass, organic dyes, doped optical fibers, and semiconductor heterojunction diodes. Note that in thermodynamic equilibrium, materials can absorb light but cannot amplify it. To achieve gain, power must be supplied to the medium to promote electrons into excited-state energy levels. When the electrons are "pumped" to the excited state at a sufficiently high rate (i.e., when enough power is supplied), a population inversion, in which the population of electrons in an excited energy level exceeds that in a lower level, can be achieved. A population inversion is a necessary condition to achieve optical gain. Power can be supplied to the laser medium in many different ways. Optical pumping, in which absorption of pump photons from a flashlamp or an external laser promotes the electrons to the excited state, is often used in ultrashort-pulse lasers. Other pumping methods include current injection in semiconductor diode lasers or electric discharges in gas lasers. The need for a pump source to obtain optical gain can be likened to the need to plug in an electronic amplifier.

To achieve steady-state laser operation, the electric field must repeat itself after a round trip through the laser cavity. Therefore, let us now consider a single round trip of a monochromatic field through a cavity. To be specific, we consider the linear cavity of Fig. 1.1a. The optical path length between mirrors is denoted l; the total optical path length is therefore 2l. The electric field amplitude of a monochromatic plane wave propagating the +z direction can be written

$$E(z,t) = \operatorname{Re}\left\{E_0 e^{j(\omega t - kz)}\right\}$$
(1.36)

where the field is now taken as a scalar for convenience. E_0 is the field amplitude just to the right of the high reflector at z = 0, ω is the angular frequency, $k = \omega n/c$ is the propagation constant, and n = n' + jn'' is the complex refractive index (n' and n'' are both real numbers). The field just before the output coupler is written

$$E = \operatorname{Re}\left\{E_0 e^{(\omega/c)n''l_g} e^{j\omega t} e^{-j(\omega/c)(n'l_g+l_a)}\right\}$$
(1.37)

In eq. (1.37) n' and n'' refer to the gain medium; the regions outside the gain medium are taken as air, with n = 1. l_g and l_a are the physical lengths of the gain medium and air region, respectively. We can now identify $l = n'l_g + l_a$ as the optical path length. We also see that for n'' > 0, the field has been amplified.

The field at z = 0, corresponding to a round-trip path through the resonator, is obtained by reflecting off the output coupler with amplitude r_2 , passing back through the cavity, and then reflecting off the high reflector with amplitude r_1 . Although ideally we would have $r_1 = 100\%$, we keep the variable r_1 in our expressions to account for any imperfections of the high reflector as well as any other losses in the laser cavity besides those arising from the output coupler. The resulting expression for the field is as follows:

$$E = \operatorname{Re}\left\{r_{1}r_{2}e^{2(\omega/c)n''l_{g}}E_{0}e^{j\omega t}e^{-j(2\omega/c)l}\right\}$$
(1.38)

To satisfy the steady-state requirement, eq. (1.38) must be equal to the initial field, eq. (1.36), evaluated at z = 0: $E = \text{Re}[E_0 e^{j\omega t}]$. This leads to two conditions: the gain condition and the phase condition. The gain condition,

$$r_1 r_2 e^{(2\omega/c)n''l_g} = 1 \tag{1.39}$$

states that the round-trip gain exactly balances the round-trip loss. The phase condition,

$$\frac{2\omega l}{c} = 2m\pi \tag{1.40}$$

requires that the round-trip phase shift is equal to an integer *m* times 2π . This means that the laser is only allowed to oscillate at certain discrete angular frequencies, given by

$$\omega_m = \frac{m\pi c}{l} \tag{1.41}$$

These frequencies are known as the *longitudinal modes* of the cavity. In terms of the frequency $f = \omega/2\pi$, the mode frequencies are

$$f_m = \frac{mc}{2l} \tag{1.42}$$

and the mode spacing (taking n' as frequency independent) is

$$\Delta f = f_m - f_{m-1} = \frac{c}{2l}$$
(1.43)



Figure 1.2 Laser gain and cavity loss spectra, longitudinal mode locations, and laser output for multimode laser operation.

A frequency-domain view of basic laser operation is pictured schematically in Fig. 1.2. A comparison of the gain vs. loss is shown at the top, with the locations of the longitudinal modes shown below. The resonator loss is taken as frequency independent, while the gain is assumed to have a bandpass spectral response. Laser oscillation occurs only for those modes where the gain lies above the loss line. In the situation shown, gain exceeds loss for several longitudinal modes, and multiple output frequencies appear simultaneously. This is called *multimode operation*.



Figure 1.3 Gain and loss spectra, longitudinal mode locations, and output of a homogeneously broadened laser under single-mode operation.

To obtain monochromatic or single-mode laser radiation, it is usually necessary to insert a frequency-dependent loss element (a filter) to ensure that gain exceeds loss for only a single longitudinal mode. This situation is sketched in Fig. 1.3. Note, however, that the steady-state laser gain condition [eq. (1.39)] requires that the gain exactly equal rather than exceed the loss. For this reason we now distinguish between small-signal gain and saturated gain. Small-signal gain is gain available under conditions of zero or at least very weak light intensity and depends only on the properties of the laser medium and the pump level. For weak spontaneous emission to build up to produce a significant laser intensity, the small-signal gain must indeed exceed the loss. However, as the intracavity intensity increases, the small-signal condition is violated. As laser photons are amplified in the laser medium through stimulated emission, at the same time electrons in the excited energy state are stimulated back down to the lower-energy state. The intracavity field extracts energy that was stored in the gain medium by the pump and reduces the number of excited-state electrons available for amplification. As a result, the gain is reduced. The actual gain that results, known as the saturated gain, depends on the properties of the laser medium, the pump level, and the intracavity laser intensity. Note that saturation phenomena are quite familiar in the context of electronic amplifiers, where the full amplifier gain is available only for input signals below a certain voltage level; for higher voltage levels the output may appear to be clipped.

To clarify the role of gain saturation in lasers, let us imagine that the pump intensity is increased slowly from zero in a single-mode laser cavity. As long as the small-signal gain remains below the loss, the laser intensity is zero. When the pump intensity is sufficient to raise the small-signal gain to equal the loss exactly, the laser reaches *threshold*. The laser power is still zero at this point. When the pump is increased above threshold, the laser power increases, and this saturates the gain. For a given pump level, the laser intensity builds up just enough to maintain the saturated gain at exactly the loss level, as pictured in Fig. 1.3. Thus, pump power above the threshold value is converted into stimulated emission. As a result, the laser intensity increases linearly with pump power above threshold.

In our discussion we have implicitly assumed that the gain medium is broadened homogeneously. This means that all the excited-state electrons have identical gain spectra, so that the overall gain is simply the gain spectrum per electron times the population difference between upper and lower laser states. Gain saturation in a homogeneously broadened medium results from a decrease in this population difference, and therefore the saturated gain has the same spectral shape as the small-signal gain.

Inhomogeneously broadened media and inhomogeneously broadened lasers, in which different excited-state electrons have different spectra, are also possible. This may result, for example, due to the fact that different impurity ions in a glass laser experience different local environments. As shown in Fig. 1.4a, the overall absorption or gain spectrum is then determined by both the individual homogeneous lines (associated with electrons with identical resonance frequencies) and the inhomogeneous distribution function $G(\omega - \Omega_0)$. When the inhomogeneous distribution is wide compared to the homogeneous linewidth, the net absorption or gain spectrum is determined mainly by $G(\omega - \Omega_0)$ and peaks at the center of the inhomogeneous distribution Ω_0 . As shown in Fig. 1.4b, the saturation behavior of an inhomogeneously broadened medium interacting with a narrow-bandwidth laser field is quite different than saturation behavior for homogeneous broadening. In particular, saturation induces a "spectral hole" in the vicinity of the laser field, while the spectrum is essentially unchanged for frequencies much more than a homogeneous linewidth away from the laser frequency. Because inhomogeneously broadened lasers are not in common use for ultrashort



Figure 1.4 (a) Absorption or gain spectrum of a medium with an inhomogeneous distribution $G(\omega - \Omega_0)$ of transition frequencies; (b) spectral hole burning associated with saturation of an inhomogeneously broadened medium by a narrowband laser field.

pulse generation, we do not discuss them further at this time. However, we return to this topic in Chapter 9, where we discuss line-shape theory and ultrafast spectroscopy techniques for probing physics associated with homogeneous and inhomogeneous broadening.

1.3.2 Gain and Gain Saturation in Four-Level Atoms

To provide further insight, we discuss gain saturation for the four-level atom. The four-level atom approximation, sketched in Fig. 1.5, is commonly used to model important mode-locking media such as Ti:sapphire or dye molecules. Here we discuss continuous-wave (CW) saturation; saturation in response to pulses is covered in Chapter 2. Note that we do not mean the word *atom* in the term *four-level atom* literally. Rather, this word refers to whatever entity (molecule, impurity complex, etc.) is active in the laser gain process.

Referring to Fig. 1.5, electrons are promoted from the lowest state, denoted level 1, up to level 2 (e.g., via optical pumping). Electrons in level 2 are usually assumed to relax very rapidly to level 3, so that the population of electrons in level 2 remains close to zero. Physically, this relaxation usually arises from vibronic rearrangement of the nuclei within the crystal-impurity or molecular system, which typically occurs on a subpicosecond time scale following the electronic transition from 1 to 2. The transition from 3 to 4 is the lasing transition, and levels 3 and 4 are the upper and lower laser levels, respectively. Electrons from level 3 can undergo stimulated emission down to level 4, giving up a photon to the laser field in the process, or they can also relax spontaneously down to level 4 with rate τ_G^{-1} , in which case their energy is not available to the laser field. τ_G is the energy storage time of the gain medium, which may range from nanoseconds to milliseconds in materials used for femtosecond pulse generation. Electrons in level 4 are again assumed to relax very rapidly



Figure 1.5 Energy-level structure for a four-level atom.

back to level 1, so that the lower laser level remains nearly empty (this is the principal advantage of four-level atoms).

Mathematically, we denote the total density of atoms (in m⁻³) as N_G , and the number density in each of the four levels as N_1 , N_2 , N_3 , and N_4 , respectively. Naturally, we have

$$N_1 + N_2 + N_3 + N_4 = N_G \tag{1.44}$$

The level populations are described by the following rate equations:

$$\dot{N}_1 = -W(N_1 - N_2) + \frac{N_4}{\tau_{41}}$$
(1.45)

$$\dot{N}_2 = W(N_1 - N_2) - \frac{N_2}{\tau_{23}}$$
 (1.46)

$$\dot{N}_3 = -S(N_3 - N_4) - \frac{N_3}{\tau_G} + \frac{N_2}{\tau_{23}}$$
 (1.47)

$$\dot{N}_4 = S(N_3 - N_4) - \frac{N_4}{\tau_{41}} + \frac{N_3}{\tau_G}$$
 (1.48)

Here *W* is the pumping rate per atom (in s^{-1}) from level 1 to 2, and its strength is controlled via the external excitation of the laser medium. For example, in the case of optical pumping, *W* is proportional to the intensity of the pump laser. *S* is the stimulated emission rate per atom (in s^{-1}), which is proportional to intensity and given in physical units by

$$S = \frac{\sigma_{34}I(\omega_{34})}{\hbar\omega_{34}} \tag{1.49}$$

Here σ_{34} is the "cross section" of the laser transition, which characterizes the strength of the laser-matter interaction, $I(\omega_{34})$ is the laser intensity inside the resonator, and $\hbar\omega_{34}$ gives the photon energy of the $3 \rightarrow 4$ transition. In the case of optical pumping, a similar expression applies for W. τ_{23} and τ_{41} are the $2 \rightarrow 3$ and $4 \rightarrow 1$ relaxation times, respectively. The gain is proportional to the population difference between the laser levels:

$$g = \frac{\sigma_{34}}{2} \left(N_3 - N_4 \right) l_g \tag{1.50}$$

The factor of 1/2 arises because we have written the gain coefficient for the *field* in eq. (1.50); this factor is not present when writing the gain coefficient for the *intensity*. In steady state all the time derivatives in the rate equations are set to zero. Furthermore, since the relaxation from levels 2 and 4 is very fast, we can also approximate $N_1 - N_2 = N_1$ and $N_3 - N_4 = N_3$ in eqs. (1.45 to 1.48). Solving for the upper-laser-level population under these conditions yields the following expression for the gain:

$$g = \left(\frac{1}{2}\right) \frac{\sigma_{34} W N_G \tau_G l_g}{1 + (W + S) \tau_G}$$
(1.51)

The gain is a function of both W and S. In the small-signal regime (S = 0), the small-signal gain g_0 is given by

$$g_0 = \left(\frac{1}{2}\right) \frac{\sigma_{34} W N_G \tau_G l_g}{1 + W \tau_G} \tag{1.52}$$

 g_0 starts at zero, then increases linearly with pump rate W at first but saturates for large W when N_3 approaches N_G . In the large-signal regime (S is finite), the gain can be rewritten

$$g = \frac{g_0}{1 + S/S_{\text{sat}}} \tag{1.53}$$

where

$$S_{\text{sat}} = W + \frac{1}{\tau_G} \tag{1.54}$$

is a saturation parameter. Within a laser, the gain first increases with increasing pump until threshold is reached. Above threshold the small-signal gain g_0 continues to increase as W increases, but the actual gain g is clamped or saturated at the value needed for threshold, which we denote g_{th} . The laser intensity (inside the laser) is found by setting eq. (1.53) to g_{th} , with the result

$$S = S_{\text{sat}} \left(\frac{g_0}{g_{\text{th}}} - 1 \right) \tag{1.55}$$

Lasing can occur in three-level systems as well as in the four-level systems discussed above. For further discussion, see standard laser texts.

1.3.3 Gaussian Beams and Transverse Laser Modes

Real laser beams have a finite transverse extent; they are not plane waves. In most cases of interest to us, however, laser beams may be considered paraxial. This means that they are made up of a superposition of plane waves with propagation vectors close to a single direction (which we take as *z*). Equivalently, the variation of the field in the transverse (x-y) direction must be much weaker than in the *z* direction. Under these conditions the electric field vector still lies mainly in the x-y plane. In the following we give a brief summary of paraxial laser beams and the related transverse mode structure of lasers. More detail can be found in most texts on lasers (e.g., [7–9]).

For a monochromatic, paraxial wave it is useful to write

$$E(z,t) = \operatorname{Re}\left\{\tilde{E}_{0}u(x, y, z)e^{j(\omega t - kz)}\right\}$$
(1.56)

where k is given by eq. (1.22) and E refers to a single transverse polarization component. In this form the most rapid wavelength-scale variation of the field is carried by the e^{-jkz}

term; u(x, y, z) is a slowly varying envelope. Substitution into the wave equation, eq. (1.15), yields

$$\nabla_T^2 u + \frac{\partial^2 u}{\partial z^2} - 2jk\frac{\partial u}{\partial z} = 0$$
(1.57)

where $\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. Since *u* is assumed to be slowly varying, $\left|\frac{\partial^2 u}{\partial z^2}\right| \ll 2k \left|\frac{\partial u}{\partial z}\right|$, and therefore $\frac{\partial^2 u}{\partial z^2}$ can be neglected. The resulting paraxial wave equation is

$$\nabla_T^2 u - 2jk\frac{\partial u}{\partial z} = 0 \tag{1.58}$$

The paraxial wave equation has Gaussian beam solutions that provide a good description of laser beams both inside and outside the laser cavity. A particularly useful solution is written as follows:

$$u_{00}(x, y, z) = \frac{w_0}{w(z)} e^{-(x^2 + y^2)/w^2(z)} e^{-jk(x^2 + y^2)/2R(z)} e^{j\phi(z)}$$
(1.59)

where we define

$$w^{2}(z) = w_{0}^{2} \left[1 + \left(\frac{z}{z_{0}}\right)^{2} \right]$$
(1.60a)

$$\frac{1}{R(z)} = \frac{z}{z^2 + z_0^2} \tag{1.60b}$$

$$\phi(z) = \tan^{-1}\left(\frac{z}{z_0}\right) \tag{1.60c}$$

$$z_0 = \frac{\pi w_0^2 n}{\lambda} \tag{1.60d}$$

This solution, sketched in Fig. 1.6, describes a beam that comes to a focus at z = 0 with a radius w_0 (measured at 1/e of the on-axis field). The beam radius w(z) changes only slowly within the region $|z| < z_0$; far outside this region the beam spreads with a half angle equal to $\lambda/\pi w_0 n$. The depth *b* over which the beam radius remains less than $\sqrt{2} w_0$ is given by $b = 2z_0$, where *b* is termed the depth of focus or the confocal parameter. R(z) is the radius of curvature of the phase fronts. $R = \infty$ at the beam waist (z = 0), corresponding to a planar phase front; for $|z| \gg z_0$, the phase fronts become spherical with a center of curvature at z = 0. Equation (1.60c) indicates a phase shift that accompanies propagation through the focal region. This Gaussian beam solution describes both laser beam propagation in free space as well as the spatial behavior of the fundamental transverse modes that are allowed inside a laser resonator with flat or spherical mirrors. The particular transverse mode of this



Figure 1.6 Gaussian beam solution of the paraxial wave equation, showing evolution of the beam radius and representative phase front.

type allowed within a given laser resonator is selected by the requirement that R(z) must match the radii of curvature of the mirrors within the laser, which determines the location of the beam waist and the value of w_0 ; for free propagation outside a laser, any solution of this type is allowed.

It is useful to know the power P carried by a Gaussian beam, which is obtained by integrating the intensity over its cross-sectional area. The result is

$$P = \int dA \frac{\varepsilon_0 nc}{2} |\tilde{E}_0|^2 |u|^2 = \frac{\varepsilon_0 nc}{2} |\tilde{E}_0|^2 \frac{w_0^2}{w^2(z)} \int_0^\infty 2\pi r \, dr \, e^{-2r^2/w^2(z)}$$
$$= \frac{1}{2} \varepsilon_0 nc |\tilde{E}_0|^2 \frac{\pi w_0^2}{2}$$
(1.61)

The power is equal to the on-axis intensity at the focus, multiplied by the area of a circle of radius $w_0/\sqrt{2}$, which is the radius of the intensity at the e^{-1} point.

Equation (1.59) is the lowest-order member of an entire family of Hermite–Gaussian solutions to the paraxial wave equation, given by

$$u_{mn}(x, y, z) \sim \frac{1}{w(z)} H_m\left(\frac{\sqrt{2} x}{w(z)}\right) H_n\left(\frac{\sqrt{2} y}{w(z)}\right) \times e^{-(x^2 + y^2)/w^2(z)} e^{-jk(x^2 + y^2)/2R(z)} e^{j(m+n+1)\phi(z)}$$
(1.62)

The $H_{\nu}(\xi)$ with ν a nonnegative integer are Hermite polynomials [10], where for example, $H_0(\xi) = 1$, $H_1(\xi) = 2\xi$, and $H_2(\xi) = 4\xi^2 - 2$. In general, the H_{ν} are polynomials of order ν , which become wider and exhibit a greater number of zero crossings as ν increases. The expressions for w(z), R(z), and $\phi(z)$ are still as given above. Although the fundamental Gaussian beam is by far the most important in most ultrafast optics applications, the higher order ($m \neq 0$ or $n \neq 0$) Hermite–Gaussian solutions will be relevant in our study for several reasons. For example, all the u_{mn} in eq. (1.62) are allowed transverse modes of laser resonators, termed TEM_{mn} modes, although the fundamental TEM₀₀ mode usually has the lower loss and is therefore favored. We shall see however that higher-order modes were significant in the discovery of the important Kerr lens mode-locking technique for short pulse generation with solid-state lasers. Additionally, the Hermite–Gaussians form a complete orthogonal basis set, which means that an arbitrary spatial beam profile can be decomposed into a superposition of these functions.

To calculate the propagation of Gaussian beams through linear optical systems, it is convenient to introduce a complex q parameter, such that

$$\frac{1}{q(z)} = \frac{1}{R(z)} - \frac{j\lambda}{\pi w^2(z)n}$$
(1.63)

The effect of the optical system can then be characterized by a bilinear transformation of q: namely,

$$q_{\rm out} = \frac{Aq_{\rm in} + B}{Cq_{\rm in} + D} \tag{1.64}$$

The coefficients appearing in eq. (1.64) are usually written in matrix form. The matrices for a few common systems are as follows:

• A length *d* of a homogeneous medium:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$$
(1.65a)

• A lens with focal length f (this matrix also applies to a spherical mirror of radius R if we set f = R/2):

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}$$
(1.65b)

• A planar interface, perpendicular to the propagation direction, from a medium with index *n*₁ to a medium with index *n*₂:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{pmatrix}$$
(1.65c)

An important property of such matrices is that the effect of cascaded systems is computed via matrix multiplication. If M_1, M_2, \ldots, M_N represent the matrices for N cascaded systems, with the beam entering system 1 first and system N last, the effect of the cascaded system is given by $M_N \cdots M_2 M_1$. For example, the matrix for a homogeneous slab of thickness d and index n and surrounded on either side by free space is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/n \end{pmatrix} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix}$$
(1.66)

The matrices given above are also useful in the ray optics description of optical systems. Here one writes

$$\begin{pmatrix} x_{\text{out}} \\ x'_{\text{out}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_{\text{in}} \\ x'_{\text{in}} \end{pmatrix}$$
(1.67)

where x_{in} and x'_{in} are the position and slope of the incoming ray, respectively, and similarly for the outgoing ray. For a given optical system, the matrixes for the ray optics description and the Gaussian beam (*q* parameter) description are identical. The ray optics description is sometimes helpful in identifying or interpreting the *ABCD* matrices. For example, the matrix in eq. (1.66) says that the slab of physical thickness *d* has an effective thickness of *d/n*. From the ray optics description, it is clear that this occurs due to bending of the rays toward the normal as they enter the medium.

The *q* parameter and the *ABCD* matrices are useful in solving for the transverse mode of laser resonators. Assume that we have calculated the matrix for one complete round trip through the resonator, starting, for example, at a particular mirror \mathcal{M} . Then one condition for an acceptable transverse mode is that the field pattern must exactly reproduce itself after one round trip, which can be formulated as

$$q = \frac{Aq + B}{Cq + D} \tag{1.68}$$

A further condition is that for a stable mode, the beam radius must be finite, which means that q must have a nonzero imaginary part. From the ray optics perspective, this stability condition means that incoming rays remain confined after many passes through the system. Multiplying through by the denominator in eq. (1.68) yields a quadratic polynomial equation for q. The root of this equation with a negative imaginary part gives q at mirror \mathcal{M} , and q in turn yields the mode size within the laser resonator. Not all resonators have stable solutions. A well-known and important example is the simple two-mirror cavity (Fig. 1.7), consisting of mirrors with radii of curvature R_1 and R_2 which are separated by distance d. After calculating the round-trip *ABCD* matrix and requiring that the roots of eq. (1.68) have nonzero imaginary part, one obtains the stability relation

$$0 < \left(1 - \frac{d}{R_1}\right) \left(1 - \frac{d}{R_2}\right) < 1 \tag{1.69a}$$

or simply

$$0 < g_1 g_2 < 1 \tag{1.69b}$$



Figure 1.7 Simple two-mirror optical resonator.

where $g_1 = 1 - d/R_1$ and $g_2 = 1 - d/R_2$ are symbols often used to express the stability condition succinctly. We make use of this stability condition when we discuss Kerr lens mode-locking in Chapter 2.

1.4 INTRODUCTION TO ULTRASHORT PULSE GENERATION THROUGH MODE-LOCKING

The single-mode laser discussed in Section 1.3 can have a very narrow optical frequency spectrum. This results in the well-known monochromaticity property of lasers. However, short pulse generation requires a broad optical bandwidth and hence multiple longitudinal mode operation. Even with the bandwidth-limiting filter removed, however, the saturated gain in a homogeneously broadened laser is below the loss level except for those modes near the peak of the gain spectrum, and this also limits the number of oscillating modes. Methods for forcing a great number of modes to oscillate to obtain broad bandwidths and ultrashort pulses are covered in detail later. In the following we analyze the time-domain characteristics of a laser which we assume already has multimode operation.

For a multimode laser we can write the electric field as follows:

$$e(z,t) = \operatorname{Re}\left\{\sum_{m} E_{m} e^{j(\omega_{m}t - k_{m}z + \phi_{m})}\right\}$$
(1.70a)

where

$$\omega_m = \omega_0 + m \,\Delta\omega = \omega_0 + \frac{2m\pi c}{L} \tag{1.70b}$$

and

$$k_m = \frac{\omega_m}{c} \tag{1.70c}$$

Henceforth we use lowercase letters [e.g., e(z, t)] for the time-domain representation of the field and capital letters (e.g., E_m) to refer to the frequency domain. Equation (1.70a) most accurately models a unidirectional ring laser, where the *z* coordinate refers to travel around the ring in the direction of laser oscillation, although the same essential results will also hold for a linear (Fabry–Perot) laser. The round-trip distance around the ring is denoted *L* (for a Fabry–Perot laser we would use L = 2l, where *l* is the Fabry–Perot mirror separation). In this formulation we are assuming that the laser is oscillating in a single transverse mode, so that the spatial profile may be dropped.

Equation (1.70a) can be rewritten as follows:

$$e(z,t) = \operatorname{Re}\left\{e^{j\omega_0(t-z/c)}\sum_m E_m e^{j\left[m\Delta\omega(t-z/c)+\phi_m\right]}\right\}$$
(1.71a)

$$= \operatorname{Re}\left\{a\left(t - \frac{z}{c}\right)e^{j\omega_0(t - z/c)}\right\}$$
(1.71b)

where

$$a\left(t-\frac{z}{c}\right) = \sum_{m} E_{m} e^{j\left[m\Delta\omega(t-z/c)+\phi_{m}\right]}$$
(1.71c)

Thus, the electric field is the product of a complex envelope function a(t - z/c) with the optical carrier $e^{j\omega_0(t-z/c)}$.

This terminology is convenient because in many cases the carrier term, which oscillates with a period of just a few femtoseconds (for visible wavelengths), varies much more rapidly than the envelope function. (However, for the shortest pulses available today, comprising just a few optical cycles, this distinction is blurred.) Both carrier and envelope functions travel around the laser cavity at the speed of light. Furthermore, for a specific location in the cavity (z fixed), the envelope function is periodic with period

$$T = \frac{2\pi}{\Delta\omega} = \frac{2l}{c} = \frac{L}{c}$$
(1.72)

This corresponds to the time required for light to make one round trip around the resonator.

If we are to say anything further about the shape of the envelope function, we need to specify the mode amplitudes E_m and phases ϕ_m . If we assume that there are N oscillating modes all with equal amplitudes E_0 and with the phases identically zero, eq. (1.71a) becomes easy to evaluate. The process (discussed later) by which the modes are held with fixed relative phases is known as *mode-locking*, and as we shall see, having all the phases equal is a particularly useful form of mode-locking. Equation (1.71a) now becomes

$$e(z,t) = \operatorname{Re}\left\{E_0 e^{j\omega_0(t-z/c)} \sum_{-(N-1)/2}^{(N-1)/2} e^{j[m\Delta\omega(t-z/c)]}\right\}$$
(1.73)

To evaluate this we substitute m' = m + (N - 1)/2 and use the summation formula

$$\sum_{m=0}^{q-1} a^m = \frac{1-a^q}{1-a} \tag{1.74}$$

Upon some further simplification, the result is

$$e(z,t) = \operatorname{Re}\left\{E_0 e^{j\omega_0(t-z/c)} \frac{\sin\left(N\,\Delta\omega/2\right)(t-z/c)}{\sin\left(\Delta\omega/2\right)(t-z/c)}\right\}$$
(1.75)

The laser intensity (I), averaged over an optical cycle, is proportional to $|e(z, t)|^2$. At a specific cavity location, say z = 0, we have

$$I(t) \sim |E_0|^2 \frac{\sin^2(N \Delta \omega t/2)}{\sin^2(\Delta \omega t/2)}.$$
 (1.76)



Figure 1.8 (a) Mode-locked laser output with constant mode phase; (b) laser output with randomly phased modes.

The resulting intensity profile is sketched in Fig. 1.8a. Some of the key points are as follows:

- The output consists of a periodic series of short pulses, with period $T = 1/\Delta f = L/c$.
- The pulse duration is approximately $\Delta t = 2\pi/N \Delta \omega = 1/N \Delta f$. Thus, the pulse duration is equal to the periodicity divided by the number of modes. Equivalently, the pulse width is equal to the inverse of the total laser bandwidth.
- The peak intensity is proportional to $N^2 |E_0|^2$. In comparison, the average intensity (averaged over the pulse period), given by the number of modes times the power per mode, is proportional to $N |E_0|^2$. Thus, the peak intensity of a mode-locked pulse is enhanced by a factor N.

A very different situation arises if we assume that the mode phases are random, as in a garden-variety multimode laser. In this case the intensity profile takes on a random appearance, as shown in Fig. 1.8b. Key points are as follows:

- The intensity fluctuates randomly about the average intensity value. The average intensity is $\sim N |E_0|^2$, which is the same as in the mode-locked case, although the peak intensity is a factor of N lower.
- The time scale over which the fluctuations vary (also known as the *correlation time*) is still $\Delta t \approx 1/N \Delta f$.
- The intensity fluctuations are still periodic with period $T = 1/\Delta f$.

In addition, in the garden-variety multimode laser, the random phases are likely to vary slowly with time as well. This introduces yet an additional degree of randomness to the multimode laser output. To avoid such fluctuations and obtain a well-defined output, one can either utilize single-mode lasers with narrow spectra, as in Section 1.3, or mode-locked lasers with narrow pulses and broad spectra.

It is also worth noting that the carrier frequency ω_0 is left arbitrary in this very simple treatment. Indeed, in many mode-locked lasers, the carrier frequency is allowed to drift. On the other hand, for some purposes it is important to stabilize and know the carrier frequency precisely. These issues are discussed in Section 7.5.

1.5 FOURIER SERIES AND FOURIER TRANSFORMS

We shall have many occasions where we wish to describe pulses in terms of a frequencydomain representation. We encountered one example in Section 1.4, where we wrote a mode-locked periodic pulse train as a summation over the longitudinal cavity modes. In this section we briefly review Fourier series and Fourier transforms, which are the mathematical tools used to convert between time- and frequency-domain representations. For more detail, see texts such as [11,12]. The key point is that any time-dependent signal can be written as a superposition or summation of sines and cosines (or complex exponentials) with different frequencies.

1.5.1 Analytical Aspects

If a function f(t) is *periodic* with period T, we can write it as a Fourier series:

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\Delta\omega t}$$
(1.77)

where $\Delta \omega = 2\pi/T$ is the [angular] frequency. The F_n are the Fourier amplitudes or Fourier coefficients and can be obtained from the time-domain signal f(t) as follows³:

$$F_n = \frac{1}{T} \int_0^T f(t) e^{-jn\Delta\omega t} dt$$
(1.78)

In the case of an aperiodic time-domain function, we can consider that the period T goes to ∞ . We then have the frequency spacing $\Delta \omega \rightarrow 0$, and therefore we replace the discrete variable $n \Delta \omega$ with a continuous variable ω . This results in the Fourier transform:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} \, d\omega \tag{1.79a}$$

and

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt$$
 (1.79b)

 $F(\omega)$ is known as the Fourier transform of f(t), and f(t) is obtained by performing the inverse Fourier transform of $F(\omega)$.

We now review a few useful properties of Fourier transforms, most of which we state without proof. For all of the following we assume that f(t) and $F(\omega)$ are a Fourier transform pair.

• *Reality condition.* If f(t) is a real function, which it will be whenever it represents an actual physical observable, $F(-\omega) = F^*(\omega)$.

³ Instead of integrating from 0 to *T*, one can get the same results by integrating over any other interval of duration *T* (e.g., -T/2 to T/2).

• Scaling formula:

If
$$h(t) = f(at)$$
, then $H(\omega) = \frac{1}{a}F\left(\frac{\omega}{a}\right)$ (1.80)

• Time-delay formula:

If
$$h(t) = f(t - \tau)$$
, then $H(\omega) = F(\omega)e^{-j\omega\tau}$ (1.81)

• Frequency-offset formula:

If
$$h(t) = f(t)e^{j\omega_0 t}$$
, then $H(\omega) = F(\omega - \omega_0)$ (1.82)

• *Convolution formula.* The convolution of two functions f(t) and g(t) is denoted f(t) * g(t) and is defined by

$$f(t) * g(t) = \int dt' f(t')g(t - t')$$

The convolution formula states that

if
$$h(t) = f(t) * g(t)$$
 then $H(\omega) = F(\omega)G(\omega)$ (1.83)

• *Parseval's theorem.* We know that the intensity I(t) of a pulse whose electric field profile is f(t) is proportional to $|f(t)|^2 = f(t)f^*(t)$. Recall that the units of intensity are W/m². The pulse energy is given by the time-integrated intensity integrated over the cross-sectional area. Parseval's theorem says that

$$\int f(t)f^*(t)\,dt = \frac{1}{2\pi}\int F(\omega)F^*(\omega)\,d\omega \tag{1.84}$$

Therefore, the time-integrated intensity is equal to the frequency-integrated intensity, except for a multiplicative factor. For this reason the quantity $|F(\omega)|^2 = F(\omega)F^*(\omega)$ is called the *power spectral density*.

The use of these formulas will be illustrated as they are needed. We do provide a few examples of the Fourier transform:

- 1. $f(t) = \delta(t)$, where $\delta(t)$ is the delta function or unit impulse function. Recall that $\delta(t) = 0$ for $t \neq 0$, $\delta(t) = \infty$ for t = 0, and $\int_{-\infty}^{\infty} dt \, \delta(t) = 1$. By substitution into eq. (1.79b), we find that $F(\omega) = 1$.
- 2. If $f(t) = e^{j\omega_0 t}$, then $F(\omega) = 2\pi\delta(\omega \omega_0)$. This is verified by plugging into eq. (1.79a).

3. f(t) is a Gaussian, $f(t) = e^{-t^2/t_p^2}$. Then

$$F(\omega) = \int dt \, e^{-t^2/t_p^2} \, e^{-j\omega t} = \int dt \, e^{-\left[(t+j\omega t_p^2/2)^2/t_p^2\right]} \, e^{-\omega^2 t_p^2/4}$$

Using the substitution $u = (t + j\omega t_p^2/2)/t_p$ and $\int_{-\infty}^{\infty} du \ e^{-u^2} = \sqrt{\pi}$, we obtain

$$F(\omega) = t_p \sqrt{\pi} \ e^{-\omega^2 t_p^2/4}$$

The procedure we have used above, called *completing the square*, is very useful for evaluating Fourier transforms of Gaussian functions.

Finally, we frequently write a pulse in terms of a slowly varying envelope function times an optical carrier term:

$$e(t) = \operatorname{Re}\left\{a(t) \ e^{j\omega_0 t}\right\} = \frac{1}{2} \left[a(t) \ e^{j\omega_0 t} + a^*(t) \ e^{-j\omega_0 t}\right]$$
(1.85)

The spectrum $E(\omega)$ is then given by

$$E(\omega) = \frac{1}{2} \left[A(\omega - \omega_0) + A^*(-\omega - \omega_0) \right]$$
(1.86a)

where

$$A(\omega) = \int_{-\infty}^{\infty} a(t)e^{-j\omega t} dt \qquad (1.86b)$$

is the Fourier transform of the envelope function a(t). Equation (1.86a) is sketched in Fig. 1.9. Separating the field into envelope and carrier terms is most useful when (as in Fig. 1.9) the bandwidth of $A(\omega)$ is much less than the carrier frequency ω_0 . In the time domain, this means that the envelope is much longer than the optical cycle.



Figure 1.9 Double-sided spectrum corresponding to eq. (1.86a).

The pulse e(t) can be obtained directly by performing the inverse Fourier transform of the double-sided spectrum described by eq. (1.86a) or from the single-sided spectrum using the formula

$$e(t) = \operatorname{Re}\left\{\frac{1}{2\pi} \int A(\omega - \omega_0) e^{j\omega t} d\omega\right\}$$
(1.87)

1.5.2 Computational Aspects

Fourier transform methods are frequently used for computations involving ultrashort pulses. Accordingly, we introduce briefly some relevant points relating to numerical computations of Fourier transforms. For detailed discussion, see [13].

Usually computations involving a continuous-time function f(t) involve only samples of f(t) at a set of discrete times t_n . Let us assume that we have N evenly spaced samples with time spacing Δ , so that $t_n = n\Delta$. We designate the sample values as $f_n = f(t_n)$ for $n = 0, 1, \dots, N - 1$. The Fourier transform of f(t) is then computed numerically as

$$F_k = \sum_{n=0}^{N-1} f_n e^{-2\pi j n k/N}$$
(1.88)

Equation (1.88) is formally called the *discrete Fourier transform*. The inverse discrete Fourier transform is given by

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{2\pi j n k/N}$$
(1.89)

Indices *n* and *k* take on integer values in the range $0, 1, \ldots, N - 1$.

Discrete Fourier transform routines are widely available in numerical computing software. Practically, it is preferred to choose the vector length N to be a power of 2. This allows the numerical package to evaluate the discrete Fourier transform using the fast Fourier transform (FFT) algorithm, which dramatically reduces computing time. Although the relationship to the continuous-time Fourier transform may appear to be straightforward, several points must be kept in mind if one is to obtain sensible computational results. These points include the following:

• Although the f_n vector contains a sequence of samples of the time-domain function f(t), it does not explicitly provide information on the sampling times t_n or the sampling interval. Similarly, the vector of Fourier amplitudes F_k does not explicitly stipulate the frequencies at which the spectrum is sampled. Hence, time and frequency vectors must be tracked by the user as auxiliary information. Note, however, that time and frequency vectors are related. If the sampling interval is known to be Δ , the time vector represents a total time span $N\Delta$; the frequency vector v_k represents a total frequency span $1/\Delta$, with $v_k = k/N\Delta$ and spacing $1/N\Delta$ between frequency samples. Here frequency refers to $v = \omega/2\pi$ and is measured in hertz (as opposed to angular frequency ω measured in rad/s).

- Inspection of eqs. (1.88) and (1.89) reveals that if indices n and k were allowed to extend beyond the stipulated range, f_n and F_k would be periodic sequences with period N. f_n and F_k for n and k in the range 0, 1, ..., N − 1 would each then represent one period of the corresponding periodic sequences. For signals that are time-limited to duration less than N∆, or frequency-limited to bandwidth less than 1/∆, meaning that the signal is nonzero only over a range of samples of length less than N, this is of little consequence. However, for signals that are not time- or frequency-limited in this way, signal content appearing at the end of the time (or frequency) vector effectively wraps around to appear at the beginning of the vector.
- Consider a continuous-time signal f(t) that is frequency-limited such that its spectrum is zero for all frequencies outside the region $-B \le v \le B$. A well-known theorem states that all information about f(t) is retained in its sampled representation provided that the sampling interval satisfies $\Delta \le 1/2B$. This requirement, called the *Nyquist criterion*, is equivalent to specifying a minimum of two samples within each period of the highest-frequency component of the frequency-limited signal. Conversely, when the Nyquist criterion is not satisfied, information is lost. In the context of our discussion on the discrete Fourier transform, we note that if F_k is frequency-limited to a bandwidth less than $1/\Delta$, we have $2B < 1/\Delta$, and the Nyquist criterion is satisfied. Accordingly, in setting up an F_k vector for a computation, it is usually advisable to ensure that the nonzero values of F_k fit comfortably within the length of the vector. If necessary, this can be accomplished by zero padding the F_k vector, which is equivalent in the time domain to decreasing the sample interval (increasing the sample rate).
- For similar reasons it is usually advisable to make sure that the nonzero values of f_n fit comfortably within the vector length as well. Again this can be accomplished by zero padding if needed.
- Inspection of eqs. (1.88) and (1.89) also shows that time t = 0 of the physical continuous-time function f(t) and frequency v = 0 of the continuous-frequency function $F(\omega)$ correspond to the first point in the f_n and F_k vectors, respectively (i.e., $t_0 = 0$ and $v_0 = 0$). For a signal with spectrum centered at v = 0, the "right" half of the spectrum appears at the beginning of the vector; the "left" half of the spectrum wraps around and appears at the end of the vector. A similar statement is true for a signal centered at t = 0 in the time domain.
- If, instead, one inadvertently positions an input spectrum so that it is centered in the frequency vector, then upon performing the inverse transform, one will notice that the expected time-domain signal is multiplied by a mysterious (-1)ⁿ function. That is, data points have alternating signs, which is equivalent to a π phase shift between points. The explanation is that the discrete Fourier transform "interprets" a shift in the positioning of the input spectrum as an actual frequency shift. Centering the input spectrum amounts to a frequency shift of 1/2∆ in real units. This corresponds to a period of 2∆ in time units, equivalent to two sample intervals.
- Similarly, if one inadvertently positions an input time-domain signal so that it is centered in the time vector, then upon performing the forward transform, one will notice that the expected spectrum is multiplied by a $(-1)^k$ function. The explanation here is that a time shift in the input signal gives rise to a sampled linear phase variation in its frequency representation (F_k). For a time shift of $N\Delta/2$, the linear spectral phase amounts to a π phase shift per frequency sample.

PROBLEMS

- 1.1. As a review of complex notation, prove eq. (1.32).
- **1.2.** Derive expressions for the saturated and small-signal gain of a three-level atom. Your analysis should parallel the analysis given in Section 1.3.2 for a four-level atom, but with level 4 missing. Assume that stimulated emission occurs between levels 3 and 1 and that the population in level 3 relaxes spontaneously to level 1 with time constant τ_G . Plot a family of curves for the saturated gain as a function of *W* (different curves represent different values of *S*) for both three- and four-level atoms. Comment on the main differences in gain behavior of three- and four-level atoms and on the implications for laser operation.
- **1.3.** Analyze the gain of a two-level atom. Now there is only one laser field, denoted W. What is the gain as a function of W and τ_G , and how does this differ from three- and four-level atoms? Explain why lasers are never based on two-level atoms.
- **1.4.** In this problem we explore the steady-state laser oscillation condition by analyzing an optical transmission resonator with gain. The setup is similar to Fig. 1.1a, but with a plane-wave field E_{in} with frequency ω incident on mirror r_1 from the left and with r_1 allowed to take on arbitrary reflectivity. The output field emerging from mirror r_2 may be written

$$E_{\text{out}} = E_{\text{in}} t_1 t_2 \sum_{n=0}^{\infty} (r_1 r_2)^n e^{(2n+1)g} e^{-(2n+1)j\omega l/c}$$

where e^g and l are the single-pass gain and optical path length, respectively, and $t_i = \sqrt{1 - r_i^2}$ is a field transmission coefficient.

- (a) Assuming zero gain (g = 0), work out a formula for the intensity transmission coefficient $|E_{out}/E_{in}|^2$. [*Hint:* Equation (1.74) may come in handy.] Assuming a symmetric resonator $(r_1 = r_2)$, plot $|E_{out}/E_{in}|^2$ as a function of ω for several values of r_1 , ranging from 90 to 99%. Comment on key features in your plots, including linewidths and trends with increasing r_1 .
- (b) Repeat part (a) but now include gain in the resonator. Plot $|E_{out}/E_{in}|^2$ vs. ω for various values of g below the laser oscillation condition. Comment on the behavior of your plots as the gain increases toward the oscillation condition. Also explain how the phase condition shows up in your plots.
- **1.5.** Verify by direction substitution that the Gaussian beam given by eq. (1.59) is a solution of the paraxial wave equation.
- **1.6.** A mode-locked laser generates pulses with 10^5 W of peak power. Spatially the laser output is a Gaussian beam. If the beam is focused in air to a diameter of 10 μ m (at e^{-2} points of the intensity), give the peak intensity and the corresponding peak electric field amplitude.

1.7. A mode-locked laser usually has a smooth frequency spectrum. Consider a Gaussian spectrum given by

$$A(\tilde{\omega}) = e^{-\tilde{\omega}^2/(\Delta\Omega)^2} \sum_m \delta(\tilde{\omega} - m \Delta\omega) e^{j\phi_m}$$

where $\tilde{\omega} = \omega - \omega_0$.

- (a) Assuming constant spectral phase, $\phi_m = 0$, work out the expression for the timedomain complex envelope function a(t). (*Hint:* The Fourier transform of a periodic train of evenly spaced, equal-amplitude delta functions is a periodic train of evenly spaced, equal-amplitude delta functions. This is shown most easily using Fourier series.)
- (b) Use a computer and an FFT algorithm to evaluate the time-domain envelope function a(t) for $\Delta\Omega/2\pi = 10$ GHz, $\Delta\omega/2\pi = 100$ MHz, and $\phi_m = 0$. Plot the spectral amplitude function $A(\tilde{\omega})$ and the temporal intensity and phase. Give the pulse duration (intensity full width at half maximum). Discuss how the setup of your array representing $A(\tilde{\omega})$ determines the number of pulses in the time-domain plot.
- (c) Now obtain the phases ϕ_m from a random number generator. Plot two examples of the temporal intensity and phase and comment on all the key differences compared to the uniform phase case.

2

PRINCIPLES OF MODE-LOCKING

2.1 PROCESSES INVOLVED IN MODE-LOCKING

The single-mode lasers reviewed in Chapter 1 were relatively simple and could be adequately modeled including only the laser gain, the cavity loss, and the effects of feedback. However, such lasers do not generate short pulses! To obtain a mode-locked laser generating ultrashort pulses, one must incorporate either active or nonlinear pulse-forming elements (modulators) into the laser cavity. Furthermore, one must also consider new processes affecting the short pulse as it propagates within the laser cavity. The following nonexhaustive list describes some of the processes that can be important in mode-locked lasers:

- *Gain.* Gain is required in mode-locked lasers just as it is in single-mode lasers. As we have seen, gain saturation, in which the gain decreases for increasing laser power, is important in any laser. In single-mode lasers, the gain is sensitive only to the average laser power. In some (not all) mode-locked lasers, however, the gain saturates dynamically on the time scale of the pulse and can be an important mode-locking process.
- *Linear loss.* Linear loss, which is also present in any laser, refers to that portion of the loss that is independent of laser power.
- *Bandwidth limitations.* Passive filtering effects can limit the bandwidth that can be transmitted through the laser cavity. Filtering can arise either from frequency-dependent loss elements or from the finite bandwidth of the laser gain medium. Bandwidth limiting effects are often undesirable in that they limit how short a pulse can be obtained.
- *Dispersion*. This refers to passive, frequency-dependent phase variations encountered by the pulse as it travels through the laser cavity. Dispersion usually leads to pulse broadening and can be an extremely important limitation for femtosecond pulse

Ultrafast Optics, By Andrew M. Weiner

Copyright © 2009 John Wiley & Sons, Inc.

generation. Consequently, dispersion compensation devices are commonly included within femtosecond laser cavities.

- Active modulation. This refers to an externally driven optical modulator that modulates either the amplitude or the phase of the circulating pulse. For mode-locking the modulation frequency is usually chosen to coincide with the cavity round-trip time. The modulator may be either an acousto-optic or electro-optic loss or phase modulator, or in some cases the gain element itself is modulated.
- *Saturable loss (self-amplitude modulation).* The cavity loss may be a function of pulse intensity or pulse energy. The loss changes dynamically in response to the pulse, which is itself modulated by the dynamically changing loss. This nonlinear process leads to a self-induced modulation with a period that is automatically equal to the cavity round-trip time. Saturable loss elements, also known as *saturable absorbers*, are classified as either fast or slow, depending on how rapidly they recover with respect to the pulse duration.
- *Self-phase modulation.* The phase can also vary nonlinearly with the time-dependent pulse intensity. This leads to a nonlinear self-phase modulation process, which again occurs at the pulse round-trip time. In mode-locked lasers the nonlinear phase shift usually has a very rapid response time and thus depends only on the instantaneous laser intensity, but in some types of lasers slower nonlinear phase processes have also been important. Self-phase modulation can interact in a very interesting and useful way with dispersion, leading to *soliton* pulses which are stable against perturbations. Such solitons are important not only in femtosecond solid-state lasers, but also in optical fiber communications.

Modulation plays a key role in initiating and maintaining mode-locked laser operation. *Active mode-locking* refers to the case where the modulator is externally driven. *Passive mode-locking* refers to situations in which the pulse forms its own modulation through nonlinearities; both amplitude and phase nonlinearities can be important. The shortest pulses have been generated using passively mode-locked lasers. In this chapter we cover the more basic concepts of active and passive mode-locking. More advanced mode-locking concepts are covered in Chapter 7 after we have provided prerequisite information on dispersion and ultrafast nonlinear optics in the necessary depth.

One can imagine many different combinations of the pulse-shaping processes listed above, corresponding to many different flavors of mode-locked lasers. In our treatment we concentrate on a few of the main possibilities that are of historical or current importance. In doing so, we will not only learn about basic mode-locking principles but also learn more about processes affecting ultrafast optical signals and about lasers in general.

2.2 ACTIVE MODE-LOCKING

We begin by discussing active mode-locking, one of the first mode-locking techniques demonstrated and conceptually one of the simplest. As sketched in Fig. 2.1a, mode-locking is achieved with the aid of an externally driven intracavity loss modulator. The modulator is driven at a frequency $\omega_m = 2\pi/T$ corresponding to the longitudinal mode spacing (where *T* is the round-trip transit time). In the time domain this means that the modulator acts like a periodic shutter which opens once per pulse round-trip time (Fig. 2.1b). We can envision



Figure 2.1 (a) Actively mode-locked laser arrangement, with an intracavity modulator driven at the cavity round-trip period; (b) periodic modulator transmission and resulting mode-locked pulses.

that the initial laser radiation is shortened again and again every time it passes through the modulator. Thus, to minimize the intracavity losses, one would expect laser emission to occur in the form of short pulses synchronized with positions of maximum shutter opening. Also in the steady state, we will find that the per-pass pulse shortening by the modulator must be balanced exactly by the (unwanted) per-pass pulse-broadening mechanisms. In the following section we analyze active mode-locking from this time-domain point of view. Subsequently, we present an alternative analysis from a frequency-domain perspective.

2.2.1 Time-Domain Treatment

In our analysis we will consider propagation of an ultrashort pulse around an actively modelocked laser cavity [8,14]. Our model for the laser cavity is shown in Fig. 2.2. The principal elements in our model are the laser cavity itself, the gain medium, and the modulator. To arrive at a stable mode-locked solution, we require that the pulse reproduce itself after one round trip through the laser.

We start off by assuming a Gaussian laser pulse of the form

$$e(t) = \operatorname{Re}\left\{a(t)e^{j\omega_0 t}\right\} = \operatorname{Re}\left\{E_0 e^{-\Gamma t^2} e^{j\omega_0 t}\right\}$$
(2.1)

where $a(t) = E_0 e^{-\Gamma t^2}$ is the envelope function and Γ is a pulse-width parameter. From our results in Section 1.5, the spectrum can be written

$$E(\omega) = \frac{1}{2} \left[A(\omega - \omega_0) + A^*(-\omega - \omega_0) \right]$$
(2.2)



Figure 2.2 Model of an actively mode-locked laser.

with

$$A(\omega) = E_0 \left(\frac{\pi}{\Gamma}\right)^{1/2} e^{-\omega^2/4\Gamma}$$

We now consider each laser element in turn.

The cavity has two effects: It introduces loss and a round-trip time delay. We can represent these effects by the multiplicative factor

$$e^{-\ell}e^{-j\omega L/c}$$

Here $e^{-\ell}$ is the round-trip attenuation of the field accounting for everything except the loss modulator. The term $e^{-j\omega L/c}$ corresponds to a delay by the round-trip propagation time L/c, where $L = l_a + n'_g l_g + n'_m l_m$ is the round-trip optical path length, l_a , l_g , and l_m are the physical lengths of air regions, gain medium, and modulator, respectively, and n'_g and n'_m are the real parts of the refractive index of the gain medium and the modulator. Note that n'_g does not include the part of the gain medium refractive index due to the laser transition (see below).

The purpose of the gain medium is, of course, to amplify the field. We make the following assumptions:

- The gain medium is homogeneously broadened and has a finite optical bandwidth. The finite-gain bandwidth acts to limit how short a pulse can be attained.
- The laser spectrum is centered at the peak of the gain spectrum.
- The relaxation time of the gain medium is much longer than the pulse repetition period. Consequently, the gain saturates only in response to the average laser power. The gain can be considered constant during the passage of a single laser pulse.

Specifically, we express the action of the gain medium by the multiplicative factor

$$\exp\left[\frac{g}{1+2j[(\omega-\omega_0)/\omega_G]}\right]$$

Here g refers to the saturated gain, and the particular gain line-shape function

$$G(\omega) = \frac{g}{1 + 2j[(\omega - \omega_0)/\omega_G]}$$
(2.3)

is known as a *Lorentzian* and corresponds to our assumption of a homogeneously broadened atomic-like gain medium. Introducing the notation $\tilde{\omega} = \omega - \omega_0$, we can rewrite eq. (2.3) as

$$G(\omega) = \frac{g\left(1 - 2j\tilde{\omega}/\omega_G\right)}{1 + \left(4\tilde{\omega}^2/\omega_G^2\right)}$$
(2.4)

36 PRINCIPLES OF MODE-LOCKING

 $G(\omega)$ can be related to the complex refractive index n = n' + jn'' by identifying $G(\omega) = -j\omega n l_g/c$ [see eq. (1.37)]. Consequently, we have

$$n' = \frac{c}{\omega l_g} \frac{2g\left(\tilde{\omega}/\omega_G\right)}{1 + \left(4\tilde{\omega}^2/\tilde{\omega}_G^2\right)}$$
(2.5a)

and

$$n'' = \frac{c}{\omega l_g} \frac{g}{1 + \left(4\tilde{\omega}^2/\tilde{\omega}_G^2\right)}$$
(2.5b)

n'' is associated with gain, and n' is associated with real refractive index (i.e., phase) effects. Figure 2.3 shows a plot of n' and n''. The gain term (n'') has a symmetric peak centered at ω_0 with a full width at half maximum (FWHM) of ω_G . n' is an antisymmetric function with a zero at ω_0 and with n' < 0 below resonance ($\omega < \omega_0$).

Before proceeding, we note that the value of the peak gain g is proportional to the population difference $(N_u - N_l)$ between upper and lower laser energy levels. For gain to occur, the population must be inverted $(N_u - N_l > 0)$, and n' and n'' are as shown in Fig. 2.3. If the pump is turned off, the population is not inverted $(N_u - N_l < 0)$, and the signs of n' and n'' are opposite to those shown in the figure. Therefore, for unpumped media, n'' < 0, signifying absorption, and n' > 0 for $\omega < \omega_0$.

In actively mode-locked lasers, the pulse usually has a narrow bandwidth compared to the gain bandwidth ω_G . Therefore, we can approximate eq. (2.3) using a second-order Taylor series expansion. Using

$$\frac{1}{1+\Delta} \approx 1 - \Delta + \Delta^2 \tag{2.6}$$



Figure 2.3 Complex refractive index plotted vs. normalized frequency for a Lorentzian gain medium. n' and n'' are normalized by the factor $cg/\omega l_g$.

which is valid for small Δ , we obtain

$$G(\omega) \approx g \left(1 - \frac{2j\tilde{\omega}}{\omega_G} - \frac{4\tilde{\omega}^2}{\omega_G^2} \right)$$
 (2.7)

Now by using eqs. (2.2) and (2.7), we can obtain an expression for the spectrum as affected by the cavity and the gain medium:

$$A(\tilde{\omega}) = E_0 \left(\frac{\pi}{\Gamma}\right)^{1/2} e^{-j\omega_0 L/c} e^{g-\ell} e^{-j\tilde{\omega}(L/c+2g/\omega_G)} e^{-\tilde{\omega}^2/4\Gamma'}$$
(2.8a)

where

$$\frac{1}{\Gamma'} = \frac{1}{\Gamma} + \frac{16g}{\omega_G^2} \tag{2.8b}$$

If the fractional change in Γ (or equivalently, in the pulse width) per pass through the gain medium is small, we can rewrite eq. (2.8b) as

$$\Gamma' \approx \Gamma - \frac{16g\Gamma^2}{\omega_G^2} \tag{2.9}$$

At this point we Fourier-transform back into the time domain. Assuming that $\omega_0 L/c = 2m\pi$ (i.e., ω_0 is equal to one of the cavity mode frequencies), we obtain the following for the envelope function a(t):

$$a(t) = E_0 e^{g-\ell} \left(\frac{\Gamma'}{\Gamma}\right)^{1/2} e^{-\Gamma' \left[t - (L/c + 2g/\omega_G)\right]^2}$$
(2.10)

We see that Γ' is decreased by passage through the gain medium. This means that the gain spectral profile reduces the bandwidth of the circulating pulse, resulting in an increase in pulse duration. Additionally, the pulse is delayed by a time

$$T = \frac{L}{c} + \frac{2g}{\omega_G} = \frac{L}{c} \left(1 + \frac{g}{\pi} \frac{\Delta\omega}{\omega_G} \right)$$
(2.11)

where $\Delta \omega = 2\pi c/L$ is the *cold cavity* (i.e., the cavity without including the effect of the laser transition) longitudinal mode spacing. The round-trip delay time is increased over that of the cold cavity (L/c) by an amount $Lg \Delta \omega/c\pi \omega_G$. This extra delay arises due to the spectral variation of n' due to the laser transition. When n' is a function of frequency, the *pulse propagation velocity* (also known as the *group velocity* v_g) differs from the *propagation velocity* of the carrier (also known as the *phase velocity* $v_p = c/n'$). We discuss group and phase velocities further in a later chapter.

We will now account for the loss modulator by multiplying eq. (2.10) by its transmission function $\eta(t)$, where

$$\eta(t) = \exp\left[-\Delta_m (1 - \cos \omega_m t)\right] \tag{2.12}$$

Thus, $\eta(t)$ varies periodically with period $2\pi/\omega_m$. For best mode-locking the modulator frequency is tuned to

$$\omega_m = \frac{2\pi}{T} \tag{2.13}$$

where *T* is the actual round-trip time given by eq. (2.11). The peak transmission is unity at t = 0, T, 2T, ..., and the minimum transmission is $e^{-2\Delta_m}$. Δ_m is called the *modulation depth*. If the pulse is much shorter than the modulation period (as indeed we desire for a mode-locked pulse!) and is centered near the times of peak transmission, we can approximate $\cos \omega_m t \approx 1 - \frac{1}{2}\omega_m^2 t^2$ near $t = 0, \cos \omega_m t \approx 1 - \frac{1}{2}\omega_m^2 (t - T)^2$ near t = T, and so on. Then the transmission for the pulse in eq. (2.10) becomes

$$\eta(t) = e^{-\frac{1}{2}\Delta_m \omega_m^2 (t-T)^2}$$
(2.14)

Multiplying (2.10) by (2.14) and using (2.11), we obtain an expression for the pulse after one round trip through the laser:

$$a(t) = E_0 e^{g-\ell} \left(\frac{\Gamma'}{\Gamma}\right)^{1/2} e^{-\Gamma''(t-T)^2}$$
(2.15a)

where

$$\Gamma'' = \Gamma' + \frac{1}{2} \Delta_m \omega_m^2 \approx \Gamma - \frac{16g\Gamma^2}{\omega_G^2} + \frac{1}{2} \Delta_m \omega_m^2$$
(2.15b)

From eq. (2.15b) the modulator decreases the pulse width.

A steady-state mode-locked solution requires that the pulse reproduce itself exactly after one round trip. Therefore, we require that eq. (2.15a) be identical to our starting point, eq. (2.1), except for the time delay. As a result, we have

$$\Gamma'' = \Gamma \tag{2.16}$$

and

$$\left(\frac{\Gamma'}{\Gamma}\right)^{1/2} e^{g-\ell} = 1 \tag{2.17}$$

Equation (2.16) says that the pulse width must reproduce itself after a round trip. Substituting in eq. (2.15b), we obtain

$$\frac{1}{2}\Delta_m \omega_m^2 - \frac{16g\Gamma^2}{\omega_G^2} = 0$$
(2.18)

and

$$\Gamma = \frac{1}{4\sqrt{2}} \left(\frac{\Delta_m}{g}\right)^{1/2} \omega_m \omega_G \tag{2.19}$$

Equation (2.18) leads to the physical interpretation that the per-pass pulse shortening due to the modulator exactly balances the per-pass broadening due to the gain medium's finite bandwidth.

Equation (2.19) is the desired solution for the pulse-width parameter Γ . It is customary to express the pulse width in terms of the full width at half maximum (FWHM) of the intensity, which we denote Δt .

$$\Delta t = 2 \left(\frac{\ln 2}{2\Gamma}\right)^{1/2} = \frac{2^{3/4} \sqrt{\ln 2}}{\pi} \left(\frac{g}{\Delta_m}\right)^{1/4} (f_m f_G)^{-1/2}$$
$$\approx 0.45 \left(\frac{g}{\Delta_m}\right)^{1/4} (f_m f_G)^{-1/2}$$
(2.20)

Here $f_m = \omega_m/2\pi$ and $f_G = \omega_G/2\pi$. In obtaining the shortest pulses, one prefers a large modulation index, a large modulation frequency, and a large gain bandwidth. The pulse width depends only weakly on these parameters, however, and usually relatively long pulses in the range 10 to 100 ps are generated.

Equation (2.17) is the gain condition for the actively mode-locked laser. The peak saturated gain (g) is slightly greater than the loss, in order to balance the losses in the modulator and the fact that the overall gain is lower than g due to the frequency spread within the mode-locked pulse and the finite gain bandwidth. The average power P circulating in the laser can be obtained from the gain saturation equation,

$$g = \frac{g_0}{1 + P/P_{\text{sat}}} \tag{2.21}$$

where g_0 is the small-signal gain and P_{sat} is the saturation power of the gain medium. In evaluating *P*, *g* can be found exactly from eq. (2.17) or can be approximated by $g \approx \ell$.

It should be noted that in our analysis, we have assumed that the modulation is exactly synchronous with the cavity transit time. Under this condition, the mode-locked pulse is positioned exactly at the peak transmission point of the applied modulation, which was also assumed. In practice, the modulation frequency must match the mode spacing very precisely, with a tolerance typically in the range of a few hundred hertz to a few kilohertz. For a typical cavity length corresponding to a repetition rate on the order of 100 MHz, this means that the frequency must be set correctly within 1 part in 10⁵ or better, and the cavity length must remain stable to better than a few tens of micrometers. When the modulation frequency is detuned (within this narrow range), the mode-locked pulse will be offset from the time of peak transmission. The nonsymmetrical variation of the modulator transmission during the duration of the mode-locked pulse will then act to push the pulse slightly back toward the transmission peak, which provides a timing restoration mechanism. Stable mode-locked operation can be maintained only when the amount of timing restoration per laser transit is enough to fully compensate the difference between the modulation period and the laser transit time.

40 PRINCIPLES OF MODE-LOCKING

2.2.2 Frequency-Domain Treatment

We now consider an alternative frequency-domain theory of active mode-locking. As before, we keep track of the effects induced by each laser element and require that the net change per round trip be zero. Here we do not assume a Gaussian pulse at the outset. Instead, by assuming that the change per element per round trip is small, we arrive at a differential equation for the mode-locked pulse. We shall follow a similar course in analyzing other types of mode-locking in later sections. A principal merit of the treatment in this section is to illustrate the locking of cavity modes leading to the term *mode-locking*. Our discussion most closely follows [15]; related treatments include [16,17].

We start with the effects of the cavity. We assume that the longitudinal modes of the cold cavity are at frequencies $\Omega_q = \omega_0 + q \Delta \omega$, where $\Delta \omega$ is the cold-cavity-mode spacing. For a low-loss (high-Q) resonator, the time variation of the field resonating with the qth cavity mode E_q can be written

$$\frac{dE_q}{dt} - j\Omega_q E_q + \frac{E_q}{\tau_0} = \text{source terms}$$
(2.22)

The identity of the source terms in our mode-locking problem will become clear soon. As before, we assume an intracavity loss modulator driven at frequency ω_m . For steady-state mode-locked operation, the laser output should be periodic with period $2\pi/\omega_m$. Hence, the laser emission can be written as a Fourier series with harmonics $\omega_n = \omega_0 + n\omega_m$, where *n* is an integer. Here we have assumed that the center frequency is still at ω_0 , which will be true when the central cavity mode exactly coincides with the peak of the gain spectrum. As long as ω_m nearly matches the mode-spacing $\Delta \omega$, each of the driven frequencies $\omega_n = \omega_0 + n\omega_m$ interacts strongly only with the corresponding cavity resonance at $\Omega_n = \omega_0 + n\Delta\omega$. The equation for the *n*th harmonic, written $E_n = \tilde{E}_n e^{j\omega_n t}$, is

$$\frac{d\tilde{E}_n}{dt} + (jn\omega_m\delta)\tilde{E}_n + \frac{\tilde{E}_n}{\tau_0} = \text{source terms}$$
(2.23a)

where

$$\delta = 1 - \frac{\Delta\omega}{\omega_m} \tag{2.23b}$$

Note also that we identify the loss per round trip, which was taken as $e^{-\ell}$ in the preceding section, with the factor e^{-T/τ_0} arising from eq. (2.22). *T* is the round-trip propagation time, as before. Thus, we can equate

$$\frac{1}{\tau_0} = \frac{\ell}{T} \tag{2.24}$$

Equivalently, the loss coefficient (τ_0^{-1}) in the rate equations for the harmonic amplitudes is obtained by taking the natural logarithm of the single-pass loss, then dividing by the round-trip time.

In a similar way, the gain is incorporated into eq. (2.23a) by adding the term $G(\omega)\tilde{E}_n/T$ to the right-hand side of the equation. Here $G(\omega)$ is approximated by eq. (2.7), as in Section 2.2.1.