

Computational Electrodynamics: The Finite-Difference Time-Domain Method

Second Edition

Allen Taflove
Susan C. Hagness



Artech House
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data

Taflove, Allen.

Computational electrodynamics : the finite-difference time-domain method.—2nd ed. /
Allen Taflove, Susan C. Hagness.

p. cm. — (Artech House antennas and propagation library)

Includes bibliographical references and index.

ISBN 1-58053-076-1 (alk. paper)

1. Electromagnetism. 2. Maxwell equations—Numerical solutions. 3. Maxwell
equations—Data processing. 4. Moments method (Statistics). 5. Integro-differential
equations—Numerical solutions. I. Hagness, Susan C. II. Title. III. Series.

QC760 .T34 2000

00-029991

537.6—dc21

CIP

British Library Cataloguing in Publication Data

Computational electrodynamics: the finite-difference
time-domain method.—2nd ed. — (Artech House antennas and
propagation library)
1. Electrodynamics—Mathematics 2. Finite differences
3. Time-domain analysis
I. Taflove, Allen II. Hagness, Susan C.
537.6'0151

ISBN 1-58053-076-1

Cover design by Igor Valdman

© 2000 ARTECH HOUSE, INC.

685 Canton Street
Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book
may be reproduced or utilized in any form or by any means, electronic or mechanical,
including photocopying, recording, or by any information storage and retrieval system,
without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have
been appropriately capitalized. Artech House cannot attest to the accuracy of this
information. Use of a term in this book should not be regarded as affecting the validity of
any trademark or service mark.

International Standard Book Number: 1-58053-076-1

Library of Congress Catalog Card Number: 00-029991

10 9 8 7 6 5 4 3 2 1

Contents

Preface to the Second Edition	xvii
Preface to the First Edition	xxi
1 Electrodynamics Entering the 21st Century	1
1.1 Introduction	1
1.2 The Heritage of Military Defense Applications	2
1.3 Frequency-Domain Solution Techniques	3
1.4 Rise of Finite-Difference Time-Domain Methods	3
1.5 History of FDTD Techniques for Maxwell's Equations	5
1.6 Characteristics of FDTD and Related Space-Grid Time-Domain Techniques	7
1.6.1 Classes of Algorithms	7
1.6.2 Predictive Dynamic Range	17
1.6.3 Scaling to Very Large Problem Sizes	18
1.7 Examples of Applications (including Color Plate Section, pages 9–16)	19
1.7.1 Radar-Guided Missile	9, 20
1.7.2 High-Speed Computer Circuit-Board Module	10, 21
1.7.3 Power-Distribution System for a High-Speed Computer Multichip Module	11, 22
1.7.4 Microwave Amplifier	12, 23
1.7.5 Cellular Telephone	13, 24
1.7.6 Optical Microdisk Resonator	14, 25
1.7.7 Photonic Bandgap Microcavity Laser	15, 27
1.7.8 Colliding Spatial Solitons	16, 28
1.8 Conclusions	29
References	30
2 The One-Dimensional Scalar Wave Equation	35
2.1 Introduction	35
2.2 Propagating-Wave Solutions	35
2.3 Dispersion Relation	36
2.4 Finite Differences	38
2.5 Finite-Difference Approximation of the Scalar Wave Equation	39
2.6 Numerical Dispersion Relation	42
2.6.1 Case 1: Very Fine Sampling in Time and Space ($\Delta t \rightarrow 0$, $\Delta x \rightarrow 0$)	43
2.6.2 Case 2: Magic Time-Step ($c\Delta t = \Delta x$)	43
2.6.3 Case 3: Dispersive Wave Propagation	44
2.6.4 Example of Calculation of Numerical Phase Velocity and Attenuation	49
2.6.5 Examples of Calculations of Pulse Propagation	51
2.7 Numerical Stability	55
2.7.1 Complex-Frequency Analysis	55
2.7.2 Examples of Calculations Involving Numerical Instability	59
2.8 Summary	61

Appendix 2A: Order of Accuracy	63
2A.1 Lax-Richtmyer Equivalence Theorem	63
2A.2 Limitations	64
References	64
Bibliography on Stability of Finite-Difference Methods	65
Problems	65
3 Introduction to Maxwell's Equations and the Yee Algorithm	67
3.1 Introduction	67
3.2 Maxwell's Equations in Three Dimensions	67
3.3 Reduction to Two Dimensions	70
3.3.1 TM_z Mode	71
3.3.2 TE_z Mode	71
3.4 Reduction to One Dimension	72
3.4.1 x -Directed, z -Polarized TEM Mode	72
3.4.2 x -Directed, y -Polarized TEM Mode	73
3.5 Equivalence to the Wave Equation in One Dimension	74
3.6 The Yee Algorithm	75
3.6.1 Basic Ideas	75
3.6.2 Finite Differences and Notation	77
3.6.3 Finite-Difference Expressions for Maxwell's Equations in Three Dimensions	80
3.6.4 Space Region With a Continuous Variation of Material Properties	85
3.6.5 Space Region With a Finite Number of Distinct Media	87
3.6.6 Space Region With Nonpermeable Media	89
3.6.7 Reduction to the Two-Dimensional TM_z and TE_z Modes	91
3.6.8 Interpretation as Faraday's and Ampere's Laws in Integral Form	93
3.6.9 Divergence-Free Nature	96
3.7 Alternative Finite-Difference Grids	98
3.7.1 Cartesian Grids	99
3.7.2 Hexagonal Grids	101
3.7.3 Tetradecahedron / Dual-Tetrahedron Mesh in Three Dimensions	104
3.8 Summary	105
References	106
Problems	106
4 Numerical Dispersion and Stability	109
4.1 Introduction	109
4.2 Derivation of the Numerical Dispersion Relation for Two-Dimensional Wave Propagation	110
4.3 Extension to Three Dimensions	112
4.4 Comparison With the Ideal Dispersion Case	113
4.5 Anisotropy of the Numerical Phase Velocity	114
4.5.1 Sample Values of Numerical Phase Velocity	114
4.5.2 Intrinsic Grid Velocity Anisotropy	120
4.6 Complex-Valued Numerical Wavenumbers	124
4.6.1 Case 1: Numerical Wave Propagation Along the Principal Lattice Axes	124
4.6.2 Case 2: Numerical Wave Propagation Along a Grid Diagonal	127
4.6.3 Example of Calculation of Numerical Phase Velocity and Attenuation	129
4.6.4 Example of Calculation of Wave Propagation	131

4.7	Numerical Stability	133
4.7.1	Complex-Frequency Analysis	133
4.7.2	Example of a Numerically Unstable Two-Dimensional FDTD Model	139
4.8	Generalized Stability Problem	141
4.8.1	Boundary Conditions	141
4.8.2	Variable and Unstructured Meshing	142
4.8.3	Lossy, Dispersive, Nonlinear, and Gain Materials	142
4.9	Modified Yee-Based Algorithms for Improved Numerical Dispersion	142
4.9.1	Strategy 1: Center a Specific Numerical Phase-Velocity Curve About c	143
4.9.2	Strategy 2: Use Fourth-Order-Accurate Spatial Differences	143
4.9.3	Strategy 3: Use Hexagonal Grids	152
4.9.4	Strategy 4: Use Discrete Fourier Transforms to Calculate the Spatial Derivatives	156
4.10	Alternating-Direction-Implicit Time-Stepping Algorithm for Operation Beyond the Courant Limit	160
4.10.1	Numerical Formulation of the Zheng/Chen/Zhang Algorithm	162
4.10.2	Numerical Stability	169
4.10.3	Numerical Dispersion	171
4.10.4	Discussion	171
4.11	Summary	172
	References	172
	Problems	173
	Projects	174

5 Incident Wave Source Conditions 175

5.1	Introduction	175
5.2	Pointwise \vec{E} and \vec{H} Hard Sources in One Dimension	176
5.3	Pointwise \vec{E} and \vec{H} Hard Sources in Two Dimensions	178
5.3.1	Green's Function for the Scalar Wave Equation in Two Dimensions	178
5.3.2	Obtaining Comparative FDTD Data	179
5.3.3	Results for Effective Action Radius of a Hard-Sourced Field Component	180
5.4	\vec{J} and \vec{M} Current Sources in Three Dimensions	182
5.4.1	Sources and Charging	183
5.4.2	Sinusoidal Sources	184
5.4.3	Transient (Pulse) Sources	185
5.4.4	Intrinsic Lattice Capacitance	189
5.4.5	Intrinsic Lattice Inductance	190
5.4.6	Impact Upon FDTD Simulations of Lumped-Element Capacitors and Inductors	191
5.5	The Plane-Wave Source Condition	193
5.6	The Total-Field / Scattered-Field Technique: Ideas and One-Dimensional Formulation	194
5.6.1	Ideas	194
5.6.2	One-Dimensional Formulation	197
5.7	Two-Dimensional Formulation of the TF / SF Technique	201
5.7.1	Consistency Conditions	203
5.7.2	Calculation of the Incident Field	207
5.7.3	Illustrative Example	212
5.8	Three-Dimensional Formulation of the TF / SF Technique	212
5.8.1	Consistency Conditions	216
5.8.2	Calculation of the Incident Field	221
5.9	Pure Scattered-Field Formulation	224

5.9.1	Application to PEC Structures	224
5.9.2	Application to Lossy Dielectric Structures	225
5.9.3	Choice of Incident Plane-Wave Formulation	227
5.10	Waveguide Source Conditions	227
5.10.1	Pulsed Electric Field Modal Hard Source	228
5.10.2	Total-Field / Reflected-Field Modal Formulation	229
5.10.3	Resistive Source and Load Conditions	230
5.11	Summary	231
	References	232
	Problems	232
	Projects	232

6 Analytical Absorbing Boundary Conditions 235

6.1	Introduction	235
6.2	Bayliss-Turkel Radiation Operators	237
6.2.1	Spherical Coordinates	238
6.2.2	Cylindrical Coordinates	241
6.3	Engquist-Majda One-Way Wave Equations	244
6.3.1	One-Term and Two-Term Taylor Series Approximations	245
6.3.2	Mur Finite-Difference Scheme	248
6.3.3	Trefethen-Halpern Generalized and Higher Order ABCs	251
6.3.4	Theoretical Reflection Coefficient Analysis	253
6.3.5	Numerical Experiments	256
6.4	Higdon Radiation Operators	261
6.4.1	Formulation	261
6.4.2	First Two Higdon Operators	263
6.4.3	Discussion	264
6.5	Liao Extrapolation in Space and Time	265
6.5.1	Formulation	265
6.5.2	Discussion	267
6.6	Ramahi Complementary Operators	269
6.6.1	Basic Idea	269
6.6.2	Complementary Operators	270
6.6.3	Effect of Multiple Wave Reflections	271
6.6.4	Basis of the Concurrent Complementary Operator Method	273
6.6.5	Illustrative FDTD Modeling Results Obtained Using the C-COM	278
6.7	Summary	281
	References	281
	Problems	282

7 Perfectly Matched Layer Absorbing Boundary Conditions 285 (Stephen D. Gedney and Allen Taflove)

7.1	Introduction	285
7.2	Plane Wave Incident Upon a Lossy Half-Space	286
7.3	Plane Wave Incident Upon Berenger's PML Medium	288
7.3.1	Two-Dimensional TE_z Case	289
7.3.2	Two-Dimensional TM_z Case	293
7.3.3	Three-Dimensional Case	294

7.4	Stretched-Coordinate Formulation of Berenger's PML	295
7.5	An Anisotropic PML Absorbing Medium	298
7.5.1	Perfectly Matched Uniaxial Medium	298
7.5.2	Relationship to Berenger's Split-Field PML	301
7.5.3	A Generalized Three-Dimensional Formulation	302
7.5.4	Inhomogeneous Media	304
7.6	Theoretical Performance of the PML	305
7.6.1	The Continuous Space	305
7.6.2	The Discrete Space	305
7.7	Efficient Implementation of UPML in FDTD	308
7.7.1	Derivation of the Finite-Difference Expressions	308
7.7.2	Computer Implementation of the UPML	311
7.8	Numerical Experiments With Berenger's Split-Field PML	314
7.8.1	Outgoing Cylindrical Wave in a Two-Dimensional Open-Region Grid	314
7.8.2	Outgoing Spherical Wave in a Three-Dimensional Open-Region Lattice	316
7.8.3	Dispersive Wave Propagation in Metal Waveguides	318
7.8.4	Dispersive and Multimode Wave Propagation in Dielectric Waveguides	320
7.9	Numerical Experiments With UPML	322
7.9.1	Current Source Radiating in an Unbounded Two-Dimensional Region	322
7.9.2	Highly Elongated Domains	327
7.9.3	Microstrip Transmission Line	330
7.10	UPML Termination for Conductive Media	332
7.10.1	Theory	332
	Numerical Example: Termination of a Conductive Half-Space Medium	335
7.11	UPML Termination for Dispersive Media	338
7.11.1	Theory	338
	Numerical Example: Reflection by a Lorentz Medium	343
7.12	Summary and Conclusions	343
	References	345
	Projects	347
8	Near-to-Far-Field Transformation	349
8.1	Introduction	349
8.2	Two-Dimensional Transformation, Phasor Domain	350
8.2.1	Application of Green's Theorem	351
8.2.2	Far-Field Limit	352
8.2.3	Reduction to Standard Form	354
8.3	Obtaining Phasor Quantities Via Discrete Fourier Transformation	356
8.4	Surface Equivalence Theorem	359
8.5	Extension to Three Dimensions, Phasor Domain	361
8.6	Time-Domain Near-to-Far-Field Transformation	366
8.7	Summary	371
	References	372
	Project	372
9	Dispersive and Nonlinear Materials	373
9.1	Introduction	373
9.2	Types of Dispersions Considered	374

9.2.1	Debye Media	374
9.2.2	Lorentz Media	375
9.3	Piecewise-Linear Recursive Convolution Method, Linear Material Case	375
9.3.1	General Formulation of the Method	375
9.3.2	Application to Debye Media	378
9.3.3	Application to Lorentz Media	379
9.3.4	Numerical Results	380
9.4	Piecewise-Linear Recursive Convolution Method, Nonlinear Dispersive Material Case	382
9.4.1	Governing Equations	382
9.4.2	General Formulation of the Method	384
9.4.3	FDTD Realization in One Dimension	386
9.4.4	Numerical Results	388
9.5	Auxiliary Differential Equation Method, Linear Material Case	392
9.5.1	Formulation for Multiple Debye Poles	392
9.5.2	Formulation for Multiple Lorentz Pole Pairs	394
9.5.3	Numerical Results	397
9.6	Auxiliary Differential Equation Method, Nonlinear Dispersive Material Case	398
9.6.1	Formulation for Multiple Lorentz Pole Pairs, TM_z Case	398
9.6.2	Numerical Results for Temporal Solitons	401
9.6.3	Numerical Results for Spatial Solitons	404
9.7	Summary and Conclusions	407
	References	408
	Problems	409
	Projects	410

10 Local Subcell Models of Fine Geometrical Features

10.1	Introduction	411
10.2	Basis of Contour-Path FDTD Modeling	412
10.3	The Simplest Contour-Path Subcell Models	413
10.3.1	Diagonal Split-Cell Model for PEC Surfaces	413
10.3.2	Average Properties Model for Material Surfaces	415
10.4	The Contour-Path Model of the Narrow Slot	416
10.5	The Contour-Path Model of the Thin Wire	420
10.6	Locally Conformal Models of Curved Surfaces	424
10.6.1	Dey-Mitra Technique for PEC Structures	424
10.6.2	Illustrative Results for PEC Structures	427
10.6.3	Dey-Mitra Technique for Material Structures	433
10.7	Maloney-Smith Technique for Thin Material Sheets	434
10.7.1	Basis	434
10.7.2	Illustrative Results	438
10.8	Dispersive Surface Impedance	442
10.8.1	Maloney-Smith Method	442
10.8.2	Beggs Method	449
10.8.3	Lee Method	457
10.9	Relativistic Motion of PEC Boundaries	461
10.9.1	Basis	461
10.9.2	Illustrative Results	465
10.10	Summary and Discussion	468
	References	470

Bibliography	471
Projects	472
11 Nonorthogonal and Unstructured Grids	473
(Stephen D. Gedney and Faiza Lansing)	
11.1 Introduction	473
11.2 Nonuniform Orthogonal Grids	474
11.3 Locally Conformal Grids, Globally Orthogonal	482
11.4 Global Curvilinear Coordinates	484
11.4.1 Nonorthogonal Curvilinear FDTD Algorithm	484
11.4.2 Stability Criterion	490
11.5 Irregular Nonorthogonal Structured Grids	493
11.6 Irregular Nonorthogonal Unstructured Grids	500
11.6.1 Generalized Yee Algorithm	500
11.6.2 Inhomogeneous Media	506
11.6.3 Practical Implementation of the Generalized Yee Algorithm	508
11.7 A Planar Generalized Yee Algorithm	509
11.7.1 Time-Stepping Expressions	510
11.7.2 Projection Operators	511
11.7.3 Efficient Time-Stepping Implementation	513
11.8 Examples of Passive-Circuit Modeling Using the Planar Generalized Yee Algorithm	514
11.8.1 32-GHz Wilkinson Power Divider	514
11.8.2 32-GHz Gysel Power Divider	517
11.8.3 Signal Lines in an IBM Thermal Conduction Module	518
11.9 Summary and Conclusions	522
References	523
Problems	525
Projects	527
12 Bodies of Revolution	529
(Thomas G. Jurgens, Jeffrey G. Blaschak, and Gregory W. Saewert)	
12.1 Introduction	529
12.2 Field Expansion	530
12.3 Difference Equations for Off-Axis Cells	530
12.3.1 Ampere's Law Contour Path Integral to Calculate e_r	532
12.3.2 Ampere's Law Contour Path Integral to Calculate e_ϕ	534
12.3.3 Ampere's Law Contour Path Integral to Calculate e_z	536
12.3.4 Difference Equations	539
12.3.5 Surface-Conforming Contour Path Integrals	542
12.4 Difference Equations for On-Axis Cells	544
12.4.1 Ampere's Law Contour Path Integral to Calculate e_z on the z-Axis	544
12.4.2 Ampere's Law Contour Path Integral to Calculate e_ϕ on the z-Axis	546
12.4.3 Faraday's Law Calculation of h_r on the z-Axis	548
12.5 Numerical Stability	549
12.6 PML Absorbing Boundary Condition	549
12.6.1 BOR-FDTD Background	549
12.6.2 Extension of PML to the General BOR Case	552
12.6.3 Examples	558

12.7	Application to Particle Accelerator Physics	560
12.7.1	Definitions and Concepts	560
12.7.2	Examples	563
12.8	Summary	566
	References	566
	Problems	567
	Projects	568

13 Analysis of Periodic Structures

(James G. Maloney and Morris P. Kesler)

569

13.1	Introduction	569
13.2	Review of Scattering From Periodic Structures	572
13.3	Direct Field Methods	575
13.3.1	Normal Incidence Case	575
13.3.2	Multiple Unit Cells for Oblique Incidence	577
13.3.3	Sine-Cosine Method	579
13.3.4	Angled-Update Method	580
13.4	Introduction to the Field-Transformation Technique	584
13.5	Multiple-Grid Approach	589
13.5.1	Formulation	589
13.5.2	Numerical Stability Analysis	591
13.5.3	Numerical Dispersion Analysis	592
13.5.4	Lossy Materials	593
13.5.5	Lossy Screen Example	595
13.6	Split-Field Method, Two Dimensions	596
13.6.1	Formulation	596
13.6.2	Numerical Stability Analysis	598
13.6.3	Numerical Dispersion Analysis	600
13.6.4	Lossy Materials	601
13.6.5	Lossy Screen Example	602
13.7	Split-Field Method, Three Dimensions	603
13.7.1	Formulation	603
13.7.2	Numerical Stability Analysis	607
13.7.3	UPML Absorbing Boundary Condition	610
13.8	Application of the Periodic FDTD Method	614
13.8.1	Photonic Bandgap Structures	614
13.8.2	Frequency-Selective Surfaces	616
13.8.3	Antenna Arrays	618
13.9	Summary and Conclusions	623
	Acknowledgments	623
	References	623
	Projects	625

14 Modeling of Antennas

(James G. Maloney, Glenn S. Smith, Eric T. Thiele, and Om P. Gandhi)

627

14.1	Introduction	627
14.2	Formulation of the Antenna Problem	628
14.2.1	Transmitting Antenna	628

14.2.2 Receiving Antenna	630
14.2.3 Symmetry	630
14.2.4 Excitation	632
14.3 Antenna Feed Models	634
14.3.1 Detailed Modeling of the Feed	634
14.3.2 Simple Gap Feed Model for a Monopole Antenna	636
14.3.3 Improved Simple Feed Model	639
14.4 Near-to-Far-Field Transformations	644
14.4.1 Use of Symmetry	645
14.4.2 Time-Domain Near-to-Far-Field Transformation	646
14.4.3 Frequency-Domain Near-Field to Far-Field Transformation	648
14.5 Plane-Wave Source	649
14.5.1 Effect of an Incremental Displacement of the Surface Currents	649
14.5.2 Effect of an Incremental Time Shift	651
14.5.3 Relation to Total-Field / Scattered-Field Lattice Zoning	652
14.6 Case Study I: The Standard-Gain Horn	653
14.7 Case Study II: The Vivaldi Slotline Array	659
14.7.1 Background	659
14.7.2 The Planar Element	660
14.7.3 The Vivaldi Pair	665
14.7.4 The Vivaldi Quad	665
14.7.5 The Linear Phased Array	668
14.7.6 Phased-Array Radiation Characteristics Indicated by the FDTD Modeling	669
14.7.7 Active Impedance of the Phased Array	669
14.8 Near-Field Simulations	675
14.8.1 Generic 900-MHz Cellphone Handset in Free Space	675
14.8.2 900-MHz Dipole Antenna Near a Layered Bone-Brain Half-Space	676
14.8.3 840-MHz Dipole Antenna Near a Rectangular Brain Phantom	678
14.8.4 900-MHz Infinitesimal Dipole Antenna Near a Spherical Brain Phantom	680
14.8.5 1,900-MHz Half-Wavelength Dipole Near a Spherical Brain Phantom	681
14.9 Selected Recent Applications	682
14.9.1 Use of Photonic-Bandgap Materials	682
14.9.2 Ground-Penetrating Radar	682
14.9.3 Antenna-Radome Interaction	687
14.9.4 Personal Wireless Communications Devices	690
14.9.5 Biomedical Applications of Antennas	696
14.10 Summary and Conclusions	697
References	697
Projects	701

15 High-Speed Electronic Circuits With Active and Nonlinear Components

(Melinda Piket-May, Bijan Houshmand, and Tatsuo Itoh)

703

15.1 Introduction	703
15.2 Basic Circuit Parameters	705
15.2.1 Transmission Line Parameters	706
15.2.2 Impedance	706
15.2.3 S Parameters	707

15.3	Differential Capacitance Calculation	708
15.4	Differential Inductance Calculation	709
15.5	Lumped Inductance Due to a Discontinuity	711
15.5.1	Flux / Current Definition	711
15.5.2	Fitting $Z(\omega)$ or $S(\omega)$ to an Equivalent Circuit	712
15.5.3	Discussion: Choice of Methods	712
15.6	Inductance of Complex Power-Distribution Systems	713
15.6.1	Method Description	713
15.6.2	Example: Multiplane Meshed Printed-Circuit Board	715
15.6.3	Discussion	718
15.7	Parallel Coplanar Microstrips	718
15.8	Multilayered Interconnect Modeling Example	720
15.9	Digital Signal Processing and Spectrum Estimation	721
15.9.1	Prony's Method	721
15.9.2	Autoregressive Models	725
15.9.3	Padé Approximation	729
15.10	Modeling of Lumped Circuit Elements	734
15.10.1	FDTD Formulation Extended to Circuit Elements	734
15.10.2	The Resistor	736
15.10.3	The Resistive Voltage Source	737
15.10.4	The Capacitor	737
15.10.5	The Inductor	738
15.10.6	The Diode	740
15.10.7	The Bipolar Junction Transistor	741
15.11	Direct Linking of FDTD and SPICE	743
15.11.1	Basic Idea	745
15.11.2	Norton Equivalent Circuit "Looking Into" the FDTD Space Lattice	745
15.11.3	Thevenin Equivalent Circuit "Looking Into" the FDTD Space Lattice	748
15.12	Case Study: A 6-GHz MESFET Amplifier Model	750
15.12.1	Large-Signal Model	750
15.12.2	Amplifier Configuration	753
15.12.3	Analysis of the Circuit Without the Packaging Structure	754
15.12.4	Analysis of the Circuit With the Packaging Structure	756
15.13	Summary and Conclusions	759
	Acknowledgments	760
	References	761
	Additional Bibliography	763
	Projects	764

16 Microcavity Optical Resonators

16.1	Introduction	767
16.2	Issues Related to FDTD Modeling of Optical Structures	768
16.2.1	Optical Waveguides	768
16.2.2	Material Dispersion and Nonlinearities	772
16.3	Macroscopic Modeling of Optical Gain Media	772
16.3.1	Theory	773
16.3.2	Validation Studies	776
16.4	Application to Vertical-Cavity Surface-Emitting Lasers	780
16.4.1	Passive Studies	782

16.4.2 Active Studies	784
16.5 Microcavities Based on Photonic Bandgap Structures, Quasi One-Dimensional Case	788
16.6 Microcavities Based on Photonic Bandgap Structures, Two-Dimensional Case	793
16.7 Microcavity Ring Resonators	797
16.7.1 FDTD Modeling Considerations	797
16.7.2 Coupling to Straight Waveguides	800
16.7.3 Coupling to Curved Waveguides	802
16.7.4 Elongated Ring Designs	804
16.7.5 Resonances	806
16.8 Microcavity Disk Resonators	810
16.8.1 Resonance Behavior	811
16.8.2 Suppression of Higher Order Radial Whispering-Gallery Modes	815
16.8.3 Additional FDTD Modeling Studies	819
16.9 Summary and Conclusions	819
References	822
Additional Bibliography	825
Projects	826
Acronyms	827
About the Authors	831
Index	839

Preface to the Second Edition

The first edition of this book was published in 1995. We are gratified with its high level of use by both the university and industrial-research communities. It is often the text in senior-year undergraduate and first-year graduate electrical engineering courses in computational electromagnetics, and is also frequently cited in refereed journal papers as a primary background reference for FDTD methods and applications.

We have had two primary goals in creating this second edition. First, we have worked to update the book's discussions of FDTD theory and applications to account for the continuing, rapid changes in these areas since 1995. This allows the professional engineer or scientist to have a convenient single-source reference concerning the latest FDTD techniques and research problems. Second, we have worked to enhance the educational content of the book from both a fundamental theoretical perspective and from the standpoint of the course instructor's ease of use.

Coverage of Advances in FDTD Theory and Numerical Algorithms

Specifically, this second edition contains a large body of new material that discusses in detail the following recent advances in FDTD theory and numerical algorithms:

- Complex-wavenumber theory, which places our understanding of numerical dispersion on rigorous grounds and further provides a complete picture of numerical wave propagation and possible attenuation in the FDTD space lattice;
- Complex-frequency theory, an alternative rigorous approach to numerical stability, which yields key insights into the nature of unstable numerical modes in the space lattice;
- Pseudospectral time-domain algorithms, which permit the spatial discretization to approach the Nyquist limit of two points per wavelength;

- Alternating-direction implicit algorithms, which yield provable, unconditional numerical stability in three dimensions, regardless of the size of the time step;
- Electric and magnetic current-source theory, which yields key insights into the nature of the intrinsic capacitance and inductance of lattice space cells;
- Complementary-operator and concurrent complementary-operator absorbing boundary conditions, which increase the effectiveness of analytical absorbing boundary conditions by orders of magnitude;
- Uniaxial perfectly matched layer absorbing boundary conditions for terminating space lattices containing general materials, including lossy, dispersive, and inhomogeneous dielectrics;
- Piecewise-linear recursive-convolution and simplified auxiliary-differential equation techniques for modeling dispersive and nonlinear dispersive materials;
- Simplified, numerically stable subcell models of diagonal and curved perfect electric conductor surfaces in the space lattice;
- Theory and algorithms for the analysis of periodic structures;
- Padé approximations for simple, rapid, accurate calculation of the resonant frequencies and quality factors of high- Q cavities and similar structures;
- Enhanced discussion of interfacing SPICE electronic circuits models with the FDTD space lattice, including Norton's and Thevenin's equivalent circuits for the lattice.

Coverage of Advances in FDTD Modeling Applications

In addition to theoretical advances, this second edition contains significant new material that discusses in detail the following recent advances in FDTD modeling applications:

- Periodic structures, including antenna arrays, frequency-selective surfaces, and photonic bandgap structures;
- Antennas, including the standard-gain horn, whips (monopoles) mounted on cellphones, radome interactions, ground-penetrating radar, and use of photonic bandgap materials to realize all-dielectric reflectors;
- High-speed electronic circuits, including a multiplane, meshed printed-circuit board feeding power to a multichip module, and a 6-GHz MESFET amplifier analyzed for both linear and nonlinear performance;

- Microcavity optical resonators, including microrings and microdisks for wavelength-division multiplexing, vertical-cavity surface-emitting lasers, and lasers based upon photonic-bandgap structures.

In all, this book provides 57 wide-ranging examples of FDTD modeling applications that:

- Cover the electromagnetic spectrum from radio frequencies to optical frequencies;
- Include the most exciting contemporary applications of electromagnetic wave engineering, ranging from the analysis and design of modern GHz-regime computers and personal wireless communications systems, to advanced photonic devices such as the world's smallest lasers.

These 57 examples serve not only to illustrate the power and beauty of FDTD modeling, but also to inform and excite the reader about the *integral role* that electromagnetic wave phenomena play in the design and operation of our society's most advanced electronics and photonics technologies.

Enhanced Educational Features

Finally, this second edition provides enhanced educational features including:

- Worked examples of FDTD modeling results contained within the text, which serve as answers to selected homework problems and projects;
- A CD-ROM containing MATLAB™ software for one-, two, and three-dimensional FDTD codes which readers can exercise to generate modeling examples of their own choosing. The two-dimensional code has a perfectly matched layer absorbing boundary condition. The CD-ROM also contains a mesh generator in executable form for conformal modeling of a two-dimensional perfectly conducting structure using the Dey-Mittra technique, and a separate FDTD solver incorporating this structure. Finally, the CD-ROM contains images of *all* of the figures and tables in the book as downloadable portable document format (PDF) files.

In our respective teaching experiences at Northwestern University and the University of Wisconsin–Madison, we have found that the second edition provides students with improved prospects for learning and eases the burden on their instructors. When used in a semester-length course (i.e., UW–Madison), there is sufficient time to cover in detail the first seven chapters plus the instructor's choice of an additional three chapters of interest in the remainder of the book. This includes time for the students to learn to write their own working FDTD software in one and two dimensions.

When used in a quarter-length course (i.e., Northwestern), there is sufficient time to cover the first seven chapters plus the instructor's choice of one additional chapter of interest in the remainder of the book. The entirety of the book can be covered in excellent detail in a two-semester course, leading to sufficient student background to begin Ph.D. research efforts in virtually the full range of current FDTD topics.

Acknowledgments

In accomplishing this major rewrite and update of the first edition, we gratefully acknowledge the authors of our invited chapters:

- Chapter 7— Prof. Stephen Gedney
- Chapter 11— Prof. Stephen Gedney and Dr. Faiza Lansing
- Chapter 12— Dr. Thomas Jurgens, Dr. Jeffrey Blaschak, and Mr. Gregory Saewert
- Chapter 13— Dr. James Maloney and Dr. Morris Kesler
- Chapter 14— Dr. James Maloney, Prof. Glenn Smith, Dr. Eric Thiele, and Prof. Om Gandhi
- Chapter 15— Prof. Melinda Piket-May, Dr. Bijan Houshmand, and Prof. Tatsuo Itoh

We also gratefully acknowledge the authors of our invited chapter sections: Prof. Eli Turkel (Chapter 2, Appendix 2A); Mr. Christopher Wagner and Prof. John Schneider (Chapter 5, Section 5.4); Dr. S. J. Yakura (Chapter 9, Section 9.4); and Mr. Geoffrey Waldschmidt (CD-ROM, Dey-Mittra software). Biographical sketches of all of the contributing authors appear following page 831.

Also acknowledged are the helpful contributions of our graduate students, including Milica Popovic and Snow Tseng at Northwestern University, and Tae-Woo Lee, Kristen Leininger, Xu Li, and Chonlarat Wichaidit at UW–Madison.

Finally, we acknowledge our respective family members who exhibited great patience and kept their good spirits while we worked long hours on this book. The first author appreciates the understanding and forbearance of his wife, Sylvia; 21-year old son, Michael; and 18-year old son, Nathan. The second author thanks her husband, Tim, for his love, friendship, and support. We may try their patience yet one more time in about five years, when fast-moving advances in FDTD theory and applications may indicate the need for a *third edition*.

Allen Taflove, Wilmette, Illinois
Susan C. Hagness, Madison, Wisconsin
April 10, 2000

Preface to the First Edition

Almost exactly twenty years ago, I submitted the first two journal papers of my research career to *IEEE Transactions on Microwave Theory and Techniques (MTT)*. These papers described my initial explorations of what later became known as the *finite-difference time-domain* (FDTD) method for Maxwell's equations. The two papers, and the Ph.D. dissertation research that they reported, grew from a graduate seminar course at Northwestern University in bioelectromagnetic hazards that Prof. Morris Brodwin had conducted in 1972. During my independent study for this seminar, I sought to obtain a model for UHF and microwave penetration into the human eye to better understand the formation of "microwave cataracts," which had been observed in a number of radar technicians during World War II. At first, there appeared to be no viable means to solve Maxwell's equations for the complex, three-dimensional biological tissue geometry represented by the eye, its surrounding muscle and fat tissues, and its embedding within the bony orbit of the skull. Analytical models were available for the small set of generic shapes for which the Helmholtz equation could be solved using the separation of variables technique. However, these simple shapes, including the half space, the layered half space, the sphere, the layered sphere, and the infinite cone, were very inadequate for modeling the tissue geometry of the eye. Further, a back-of-the-envelope estimate of the implications of the required tissue and wavelength space resolution ruled out the frequency-domain method of moments, then capable of solving for only a few hundred field unknowns. The eye geometry required the solution of almost 100,000 field unknowns, *two to three orders of magnitude* beyond anything published in the moment method community at that time.

Almost having given up on Prof. Brodwin's seminar project, I found myself randomly leafing through back issues of *IEEE Transactions on Antennas and Propagation* in the dark, claustrophobic stacks of the old Tech Library (replaced years later by a large, bright, open building). It was then that I stumbled upon Kane Yee's 1966 paper. Six years had gone by since its publication, with very sparse references to it recorded in *Science Citation Index*. And yet the paper was the *Grail*. I sensed that

the Yee algorithm had tremendous promise. It could handle material inhomogeneities and did not require matrix inversion, meaning that I could use the university's Control Data CDC 6400 computer to crunch the problem to its conclusion. Of course, a few "minor" details had to be solved, such as sourcing a plane wave, obtaining a rudimentary absorbing boundary condition, understanding the algorithm's numerical stability properties, and progressing from one to two to three dimensions in my code development.

So, my 1975 papers in *MTT* were published. And landed with a thud. Being eager, brash, and absolutely naive, I had expected the electromagnetics community to seize upon the marvelous Yee algorithm and apply it to *everything*. However, with the exception of a few industrial research firms and U.S. government agencies active in the electromagnetic pulse area, FDTD remained essentially unused for more than a decade.

Now, after much hard work and the welcome help of a rapidly growing user community, FDTD is being used worldwide. And for just about *everything*, as I had foreseen in 1975.

It is my profound hope that the readers of this book, whether university students or professionals, can use and enjoy its material at a number of different levels. There is sufficient tutorial exposition for a class in computational electromagnetics at the senior undergraduate or first-year graduate level, at which the students have the traditional core electromagnetics background. I have taught such a course from drafts of this book for six years at Northwestern, and similar courses have been taught for the past two years at the University of Colorado at Boulder by my colleague, Prof. Melinda Piket-May. Over a ten-week academic quarter, it is possible for students to assimilate the first eight chapters, write their own working FDTD codes in two dimensions with wave source and absorbing boundary conditions, and then use their codes to implement radiation and scattering models of their own design. Over an academic semester, it is possible to augment these basics with two or three of the remaining chapters, as selected by the instructor, to specialize in advanced topics involving unstructured meshing, electronics modeling, antennas, or parallel-processing software.

For the professional, this book is intended to be comprehensive and self-teaching. There is exposure to virtually all of the latest topics in FDTD theory and practice, and a quite exhaustive list of references and bibliographic materials.

I gratefully acknowledge the marvelous technical collaboration and warm personal friendship of Prof. Korada Umashankar in this work. I also thank the contributors of the invited chapters in this book: Prof. Stephen Gedney and Dr. Faiza Lansing, Drs. Thomas Jurgens and Gregory Saewert, Prof. Melinda Piket-May, Dr. Eric Thiele, and Dr. Stephen Barnard. Also acknowledged are the contributions of my graduate students, who did the really hard work. A special thanks goes to my steadfast friend, Evans Harrigan, who has *believed* in FDTD modeling ever since we met, and saw to it that my students had all of the time on Cray's marvelous supercomputers that they needed to do their work.

Finally, I acknowledge my wife, Sylvia, and sons, Mike and Nate, who somehow were able to keep their composure while sharing a home with a very driven person who was utterly thrilled with the highs and quite agonized with the lows of FDTD developments over twenty years.

This book is the culmination of a major phase in my life. However, the FDTD story is only beginning. Let's move on to develop detailed FDTD electromagnetics models of microchips, microlasers, and microcells, and bring the power of Maxwell's equations to bear upon society's needs in ultrahigh-speed communications technology. In this manner, electromagnetic wave specialists can augment their current role in enabling people to freely communicate with each other worldwide, at the speed of light.

Allen Taflove
Wilmette, Illinois
March 16, 1995

Chapter 1

Electrodynamics Entering the 21st Century

1.1 INTRODUCTION

Maxwell's partial differential equations of electrodynamics, formulated circa 1870, represent a fundamental unification of electric and magnetic fields predicting electromagnetic wave phenomena which Nobel Laureate Richard Feynman has called the most outstanding achievement of 19th-century science. Now engineers and scientists worldwide use computers ranging from simple desktop machines to massively parallel arrays of processors to obtain solutions of these equations for the purpose of investigating electromagnetic wave guiding, radiation, and scattering phenomena and technologies. As we begin the 21st century, it may seem a little odd to devote so much effort to solving the 19th century's best equations. Thus we ask the question: "Of what relevance are solutions of Maxwell's equations to our modern society?"

The goal of this chapter is to help answer this question. We shall discuss prospects for using numerical solutions of Maxwell's equations, in particular the *finite-difference time-domain* (FDTD) method, to help innovate and design key electrical engineering technologies ranging from cellphones and computers to lasers and photonic circuits. Whereas large-scale solutions of Maxwell's equations have been motivated in the past primarily by the requirements of military defense, as discussed below, the entire field of computational electrodynamics is shifting rapidly toward important commercial applications in high-speed communications and computing that will touch everyone in their daily lives. Ultimately, this will favorably impact the economic well-being of nations as well as their military security. It is in this context of significant electrical engineering technology advances resulting from the ability to accurately and rapidly solve Maxwell's equations on large scales that this chapter is titled "Electrodynamics Entering the 21st Century."

1.2 THE HERITAGE OF MILITARY DEFENSE APPLICATIONS

From the onset of World War II until about 1990, the answer to our question about the societal relevance of solutions of Maxwell's equations would probably have been, "We *must* have strong military defense." The development of UHF and microwave radar technology during World War II motivated early work, which proved crucial to the survival of England during the grim early days of the Battle of Britain, and subsequently to the final victory of the Allied forces. During the 45 years of Cold War that followed, the advanced development of radar remained of paramount importance as both the East and West alliances developed enormous nuclear arsenals on hair-trigger alert. Radar technologies aimed at the early warning of aircraft and missiles were subsequently met with countermeasures aimed at evading or spoofing radar detection. These were in turn met by counter-countermeasures, and so forth.

Radar encompasses a wide range of needs in solving Maxwell's equations. At the radar site, microwave sources, circuits, waveguides, and antennas must be designed to generate, transport, radiate, receive, and process electromagnetic waves. For a foe determined to press an attack despite the operation of a defensive radar system, there is the need to understand the scattering of electromagnetic waves by complex, electrically large material structures. Such understanding leads directly to materials and structure-shaping technologies for designing stealthy aircraft and missiles having reduced or confusing scattering responses.

An additional military need motivating the solution of Maxwell's equations emerged after about 1960 when it became clear that a nuclear bomb detonated above the earth's atmosphere could generate a high-level *electromagnetic pulse*, or EMP. EMP can be sufficiently intense to burn out electrical and electronic equipment on the earth's surface located hundreds of miles away from the point directly below the detonation. Equipment failures on this geographical scale could leave a nation largely defenseless against subsequent attack. Therefore, substantial efforts were devoted by the defense community to "harden" key systems to reduce their vulnerability to EMP. Here, Maxwell's equations solution technologies were aimed at predicting the level of EMP penetration and coupling into potentially vulnerable equipment, and developing cost-effective means to reduce such coupling to well below the danger point.

A third area motivating the military development of Maxwell's equations solutions was explored intensively after about 1980, when technology developments permitted the generation of steerable, *high-power microwave* (HPM) beams. In principle, such beams could neutralize electronics in the manner of EMP, but be applied on a much more selective basis for either tactical or strategic applications. On the offensive side, Maxwell's equations solutions were used to design HPM sources, circuits, and antennas to generate, transport, and radiate electromagnetic waves having intensities sufficient to ionize the air at normal pressures and temperatures. Maxwell's equations were also used to understand electromagnetic wave penetration and coupling mechanisms into potential targets of HPM, and means to mitigate these mechanisms.

1.3 FREQUENCY-DOMAIN SOLUTION TECHNIQUES

During the era reviewed above, the modeling of electromagnetic engineering systems was primarily implemented using solution techniques for the sinusoidal steady-state Maxwell's equations. Before about 1960, the principal approaches in this area involved closed-form and infinite-series analytical solutions, with numerical results from these analyses obtained using mechanical calculators. After 1960, the increasing availability of programmable electronic digital computers permitted such frequency-domain approaches to rise markedly in sophistication. Researchers were able to take advantage of the capabilities afforded by powerful new high-level programming languages such as Fortran, rapid random-access storage of large arrays of numbers, and computational speeds orders of magnitude faster than possible with mechanical calculators. In this period, the principal computational approaches for Maxwell's equations included high-frequency asymptotic methods [1, 2] and integral equations [3, 4].

However, these frequency-domain techniques have difficulties and tradeoffs. For example, while asymptotic analyses are well suited for modeling the scattering properties of electrically large complex shapes, such analyses have difficulty treating nonmetallic material composition and volumetric complexity of a structure. While integral equation methods can deal with material and structural complexity, their need to construct and solve systems of linear equations limits the electrical size of possible models, especially those requiring detailed treatment of geometric details within a volume, as opposed to just the surface shape.

While significant progress has been made in solving the ultra-large systems of equations generated by frequency-domain integral equations [5], the capabilities of even the latest such technologies are exhausted by many volumetrically complex structures of engineering interest. This also holds for frequency-domain finite-element techniques, which generate sparse rather than dense matrices. Further, the very difficult incorporation of material and device nonlinearities into frequency-domain solutions of Maxwell's equations poses a significant problem as engineers seek to design active electromagnetic / electronic and electromagnetic / quantum-optical systems such as high-speed digital circuits, microwave and millimeter-wave amplifiers, and lasers.

1.4 RISE OF FINITE-DIFFERENCE TIME-DOMAIN METHODS

During the 1970s and 1980s, several defense agencies working in the areas summarized in Section 1.2 realized the limitations of frequency-domain integral-equation solutions of Maxwell's equations. This led to early explorations of a novel alternative approach: direct time-domain solutions of Maxwell's differential (curl) equations on spatial grids or lattices. The FDTD method, introduced by Yee in 1966 [6], was the first technique in this class, and has remained the subject of continuous development. Since about 1990, when engineers in the general electromagnetics community became aware of the modeling capabilities afforded by FDTD and related techniques, the interest in this area has expanded well beyond defense technology.

There are seven primary reasons for the expansion of interest in FDTD and related computational solution approaches for Maxwell's equations:

1. *FDTD uses no linear algebra.* Being a fully explicit computation, FDTD avoids the difficulties with linear algebra that limit the size of frequency-domain integral-equation and finite-element electromagnetics models to generally fewer than 10^6 electromagnetic field unknowns. FDTD models with as many as 10^9 field unknowns have been run; there is no intrinsic upper bound to this number.
2. *FDTD is accurate and robust.* The sources of error in FDTD calculations are well understood, and can be bounded to permit accurate models for a very large variety of electromagnetic wave interaction problems.
3. *FDTD treats impulsive behavior naturally.* Being a time-domain technique, FDTD directly calculates the impulse response of an electromagnetic system. Therefore, a single FDTD simulation can provide either ultrawideband temporal waveforms or the sinusoidal steady-state response at any frequency within the excitation spectrum.
4. *FDTD treats nonlinear behavior naturally.* Being a time-domain technique, FDTD directly calculates the nonlinear response of an electromagnetic system.
5. *FDTD is a systematic approach.* With FDTD, specifying a new structure to be modeled is reduced to a problem of mesh generation rather than the potentially complex reformulation of an integral equation. For example, FDTD requires no calculation of structure-dependent Green's functions.
6. *Computer memory capacities are increasing rapidly.* While this trend positively influences all numerical techniques, it is of particular advantage to FDTD methods which are founded on discretizing space over a volume, and therefore inherently require a large random access memory.
7. *Computer visualization capabilities are increasing rapidly.* While this trend positively influences all numerical techniques, it is of particular advantage to FDTD methods which generate time-marched arrays of field quantities suitable for use in color videos to illustrate the field dynamics.

An indication of the expanding level of interest in FDTD Maxwell's equations' solvers is the hundreds of papers currently published in this area worldwide each year, as opposed to fewer than ten as recently as 1985 [7]. This expansion continues as engineers and scientists in nontraditional electromagnetics-related areas, such as digital systems and integrated optics, become aware of the power of such direct-solution techniques for Maxwell's equations.

1.5 HISTORY OF FDTD TECHNIQUES FOR MAXWELL'S EQUATIONS

We can begin to develop an appreciation of the basis, technical development, and possible future of FDTD numerical techniques for Maxwell's equations by first considering their history. The following table lists some of the key publications in this area, starting with Yee's seminal paper [6]. The reader is referred to [7] for a comprehensive bibliography of FDTD publications through 1998.

TABLE 1.1
Partial History of FDTD Techniques for Maxwell's Equations

1966	Yee [6] described the basis of the FDTD numerical technique for solving Maxwell's curl equations directly in the time domain on a space grid.
1975	Taflove and Brodin obtained the correct numerical stability criterion for Yee's algorithm, and the first sinusoidal steady-state FDTD solutions of two- and three-dimensional electromagnetic wave interactions with material structures [8, 9].
1977	Holland and Kunz and Lee applied Yee's algorithm to EMP problems [10, 11].
1980	Taflove coined the FDTD acronym and published the first validated FDTD models of sinusoidal steady-state electromagnetic wave penetration into a three-dimensional metal cavity [12].
1981	Mur published the first numerically stable, second-order accurate, absorbing boundary condition (ABC) for Yee's grid [13].
1982, 1983	Taflove and Umashankar developed the first FDTD electromagnetic wave-scattering models computing sinusoidal steady-state near-fields, far-fields, and radar cross-section for two- and three-dimensional structures [14, 15].
1986	Choi and Hoefer published the first FDTD simulation of waveguide structures [16].
1987, 1988	Kriegsmann et al. and Moore et al. published the first articles on ABC theory in <i>IEEE Trans. Antennas and Propagation</i> [17, 18].
1987, 1988, 1992	Contour-path subcell techniques were introduced by Umashankar et al. to permit FDTD modeling of thin wires and wire bundles [19], by Taflove et al. to model penetration through cracks in conducting screens [20], and by Jurgens et al. to conformally model the surface of a smoothly curved scatterer [21].
1987, 1990	Finite-element time-domain (FETD), finite-volume time-domain (FVTD), and partially or completely unstructured meshes for Maxwell's equations were introduced by Cangellaris et al. [22], Shankar et al. [23], and Madsen and Ziolkowski [24].
1988	Sullivan et al. published the first three-dimensional FDTD model of sinusoidal steady-state electromagnetic wave absorption by a complete human body [25].
1988	FDTD modeling of microstrips was introduced by Zhang et al. [26].
1990, 1991	FDTD modeling of frequency-dependent dielectric permittivity was introduced by Kashiwa and Fukai [27], Luebers et al. [28], and Joseph et al. [29].

6 Computational Electrodynamics: The Finite-Difference Time-Domain Method

- 1990, FDTD modeling of antennas was introduced by Maloney et al. [30], Katz et al. [31], and Tirkas and Balanis [32].
- 1990 FDTD modeling of picosecond optoelectronic switches was introduced by Sano and Shibata [33] and El-Ghazaly et al. [34].
- 1991– 1994 FDTD modeling of the propagation of optical pulses in nonlinear dispersive media was introduced, including the first temporal solitons in one dimension by Goorjian and Taflove [35]; studies of beam self-focusing by Ziolkowski and Judkins [36]; the first temporal solitons in two dimensions by Joseph et al. [37]; and the first spatial solitons in two dimensions by Joseph and Taflove [38].
- 1992 FDTD modeling of lumped electronic circuit elements in two dimensions was introduced by Sui et al. [39].
- 1993 Toland et al. published the first FDTD models of gain devices (tunnel diodes and Gunn diodes) exciting cavities and antennas [40].
- 1994 Thomas et al. [41] introduced a Norton's equivalent circuit for the FDTD space lattice which permits the SPICE circuit analysis tool to implement accurate sub-grid models of nonlinear electronic components or complete circuits embedded within the lattice.
- 1994 Berenger introduced the highly effective, perfectly matched layer (PML) ABC for two-dimensional FDTD grids [42], which was extended to three dimensions by Katz et al. [43], and to dispersive waveguide terminations by Reuter et al. [44].
- 1995 Gedney and Lansing introduced the planar-generalized Yee algorithm [45] which permits efficient unstructured-grid FDTD modeling of microwave and digital circuits.
- 1995, 1996 Sacks et al. [46] and Gedney [47] introduced a physically realizable, uniaxial anisotropic, perfectly matched layer (UPML) ABC.
- 1996 Krumpholz and Katehi [48] introduced the multi-resolution time-domain (MRTD) technique based upon the use of wavelet expansion functions.
- 1996, 1997 Liu [49, 50] introduced the pseudospectral time-domain (PSTD) method, which permits extremely coarse spatial sampling of the electromagnetic field at the Nyquist limit of only two points per wavelength.
- 1997, 1998 Ramahi [51, 52] introduced the methods of complementary operators and concurrent complementary operators (COM, CCOM) to implement highly effective ABCs.
- 1997, 1998 Dey and Mittra [53, 54] introduced a very simple, stable, and accurate contour-path technique to model curved metal surfaces in FDTD space lattices.
- 1998 Maloney and Kesler [55] introduced several novel means to analyze periodic structures in the FDTD space lattice.
- 1999 Schneider and Wagner [56] introduced a comprehensive analysis of FDTD grid dispersion based upon complex wavenumbers.
- 1999 Painter et al. [57] used FDTD to design, construct, and successfully test the world's smallest microcavity laser based upon a two-dimensional photonic bandgap structure.
- 2000 Zheng, Chen, and Zhang [58, 59] introduced the first three-dimensional alternating-direction implicit (ADI) FDTD algorithm with provable unconditional numerical stability regardless of the size of the time step.

1.6 CHARACTERISTICS OF FDTD AND RELATED SPACE-GRID TIME-DOMAIN TECHNIQUES

FDTD and related space-grid time-domain techniques are direct solution methods for Maxwell's curl equations. These methods employ no potentials. Rather, they are based upon volumetric sampling of the unknown electric field \vec{E} and magnetic field \vec{H} within and surrounding the structure of interest, and over a period of time. The sampling in space is at sub-wavelength (sub- λ_0) resolution set by the user to properly sample the highest near-field spatial frequencies thought to be important in the physics of the problem. Typically, 10–20 samples per λ_0 are needed. The sampling in time is selected to ensure numerical stability of the algorithm.

Overall, FDTD and related techniques are marching-in-time procedures that simulate the continuous actual electromagnetic waves in a finite spatial region by sampled-data numerical analogs propagating in a computer data space. Time-stepping continues as the numerical wave analogs propagate in the space lattice to causally connect the physics of the modeled region. For simulations where the modeled region must extend to infinity, ABCs are employed at the outer lattice truncation planes which ideally permit all outgoing wave analogs to exit the region with negligible reflection. Phenomena such as induction of surface currents, scattering and multiple scattering, aperture penetration, and cavity excitation are modeled time-step by time-step by the action of the numerical analog to the curl equations. Self-consistency of these modeled phenomena is generally assured if their spatial and temporal variations are well resolved by the space and time sampling process. In fact, the goal is to provide a self-consistent model of the mutual coupling of all of the electrically small volume cells constituting the structure and its near field, even if the structure spans tens of λ_0 in three dimensions and there are hundreds of millions of space cells.

Time-stepping is continued until the desired late-time pulse response is observed at the field points of interest. For linear wave interaction problems, the sinusoidal response at these field points can be obtained over a wide band of frequencies by discrete Fourier transformation of the computed field-versus-time waveforms at these points. Prolonged "ringing" of the computed field waveforms due to a high Q-factor or large electrical size of the structure being modeled requires a combination of extending the computational window in time and extrapolation of the windowed data before Fourier transformation.

1.6.1 Classes of Algorithms

Current FDTD and related space-grid time-domain algorithms are fully explicit solvers employing highly vectorizable and parallel schemes for time-marching the six components of \vec{E} and \vec{H} at each of the space cells. The explicit nature of the solvers is usually maintained by employing a leapfrog time-stepping scheme. Current methods differ primarily in how the space lattice is set up. In fact, gridding methods can be categorized according to the degree of structure or regularity in the mesh cells:

1. *Almost completely structured.* In this case, the space lattice is organized so that its unit cells are congruent wherever possible. The most basic example of such a mesh is the pioneering work of Yee [6], who employed a uniform Cartesian grid having rectangular cells. Staircasing was used to approximate the surface of structural features not parallel to the grid coordinate axes. Later work showed that it is possible to modify the size and shape of the space cells located immediately adjacent to a structural feature to conformally fit its surface [21, 53, 54]. This is accurate and computationally efficient for large structures because the number of modified cells is proportional to the surface area of the structure. Thus, the number of modified cells becomes progressively smaller relative to the number of regular cells filling the structure volume as its size increases. As a result, the computer resources needed to implement a fully conformal model approximate those required for a staircased model. However, a key disadvantage of this technique is that special mesh-generation software must be constructed.

2. *Surface-fitted.* In this case, the space lattice is globally distorted to fit the shape of the structure of interest. The lattice can be divided into multiple zones to accommodate a set of distinct surface features [23]. The major advantage of this approach is that well-developed mesh-generation software of this type is available. The major disadvantage is that, relative to the Yee algorithm, there is substantial added computer burden due to:

- (a) Memory allocations for the position and stretching factors of each cell;
- (b) Extra computer operations to implement Maxwell's equations at each cell and to enforce field continuity at the interfaces of adjacent cells.

Another disadvantage is the possible presence of numerical dissipation in the time-stepping algorithm used for such meshes. This can limit the range of electrical size of the structure being modeled due to numerical wave-attenuation artifacts.

3. *Completely unstructured.* In this case, the space containing the structure of interest is completely filled with a collection of lattice cells of varying sizes and shapes, but conforming to the structure surface [24]. As for the case of surface-fitted lattices, mesh-generation software is available and capable of modeling complicated three-dimensional shapes possibly having volumetric inhomogeneities. A key disadvantage of this approach is its potential for numerical inaccuracy and instability due to the unwanted generation of highly skewed space cells at random points within the lattice. A second disadvantage is the difficulty in mapping the unstructured mesh computations onto the architecture of either parallel vector computers or massively parallel machines. The structure-specific irregularity of the mesh mandates a robust pre-processing algorithm that optimally assigns specific mesh cells to specific processors.

At present, the best choice of computational algorithm and mesh remains unclear. For the next several years, we expect continued progress in this area as various groups develop their favored approaches and perform validations.

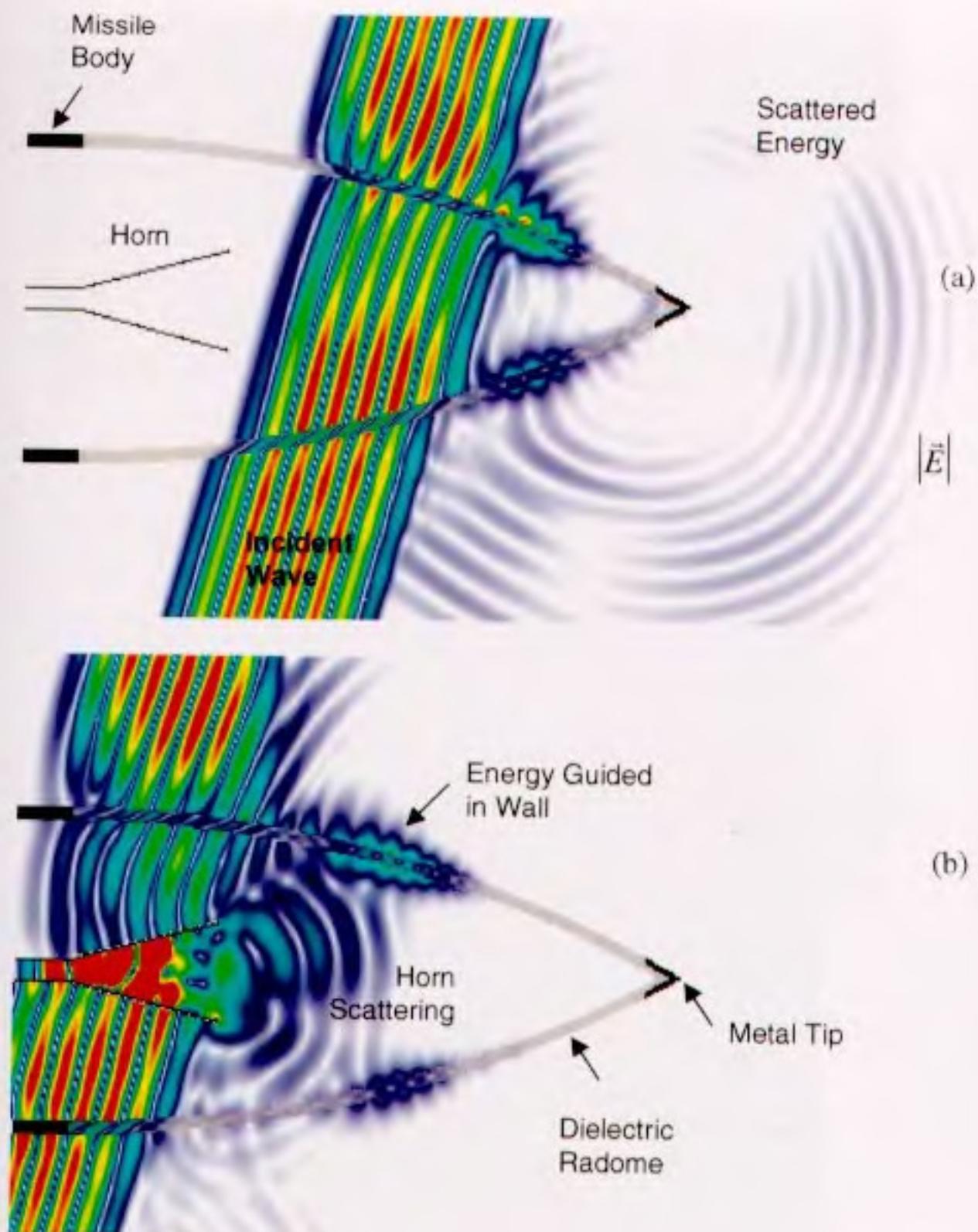


Fig. 1.1 Visualization of plane-wave penetration and scattering for the FDTD model of a missile radome containing a horn antenna. The impinging plane wave propagates from right to left and is obliquely incident at 15° from boresight. (See also Fig. 14.37.)

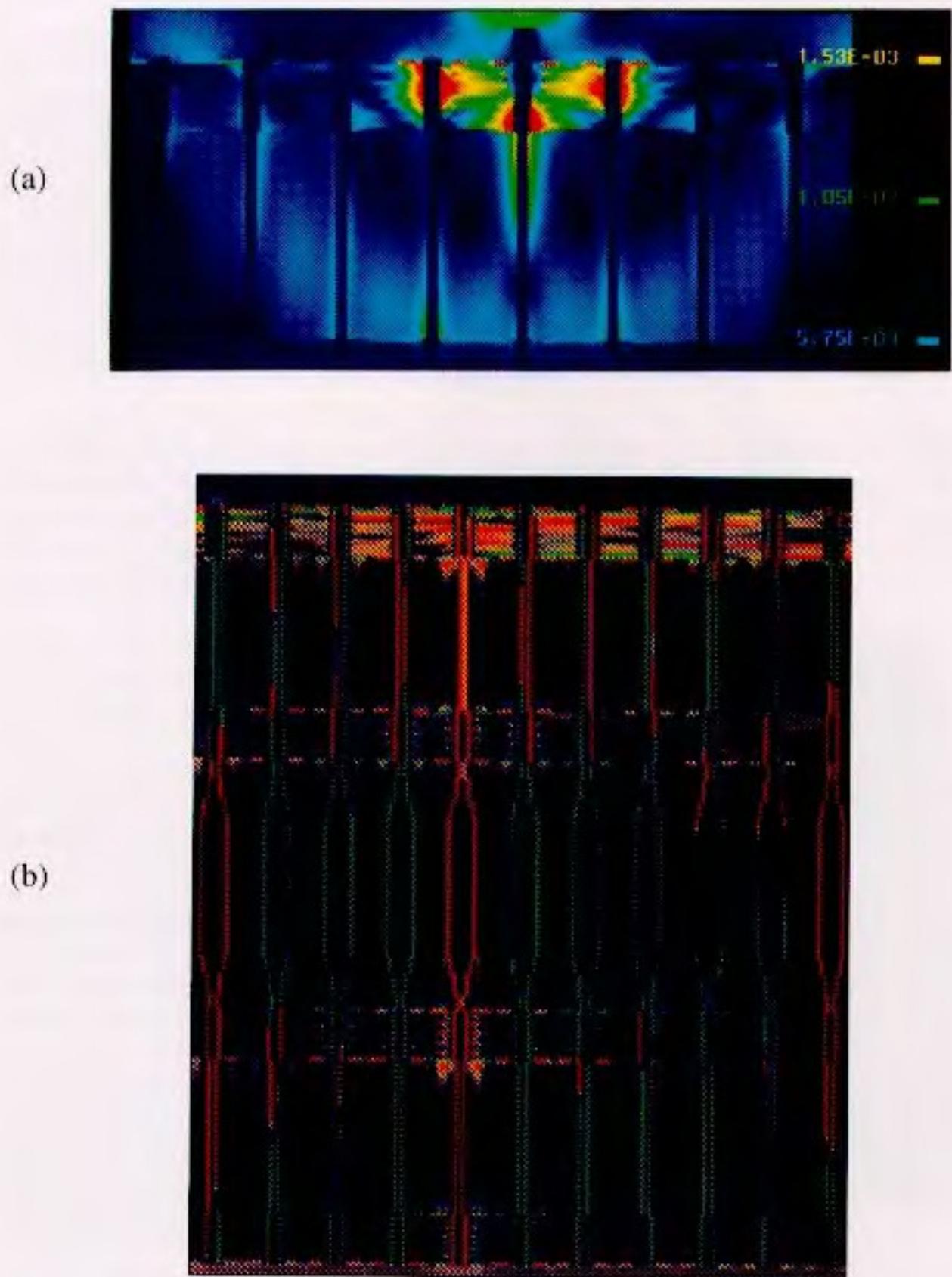


Fig. 1.2 Visualization of two modes of parasitic coupling between signal paths in a high-speed computer circuit-board module: (a) early-time magnetic fields; (b) late-time ground loops.

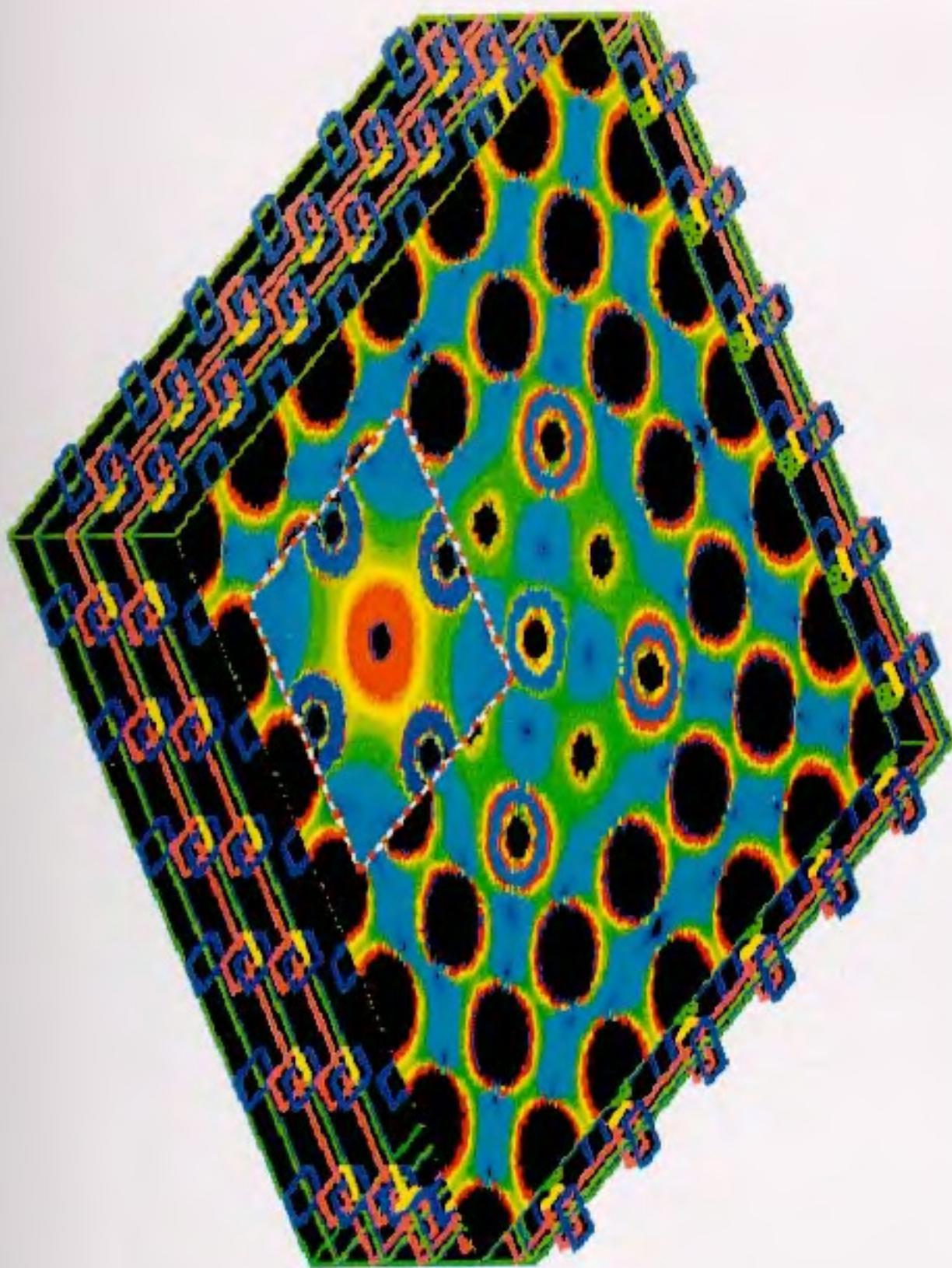


Fig. 1.3 Visualization of the FDTD-computed current density for an 8-layer power-distribution structure feeding a high-speed multichip (MCM). The MCM ground layer is located within the dotted rectangle of the meshed printed-circuit ground plane. (See also Fig. 15.4.)

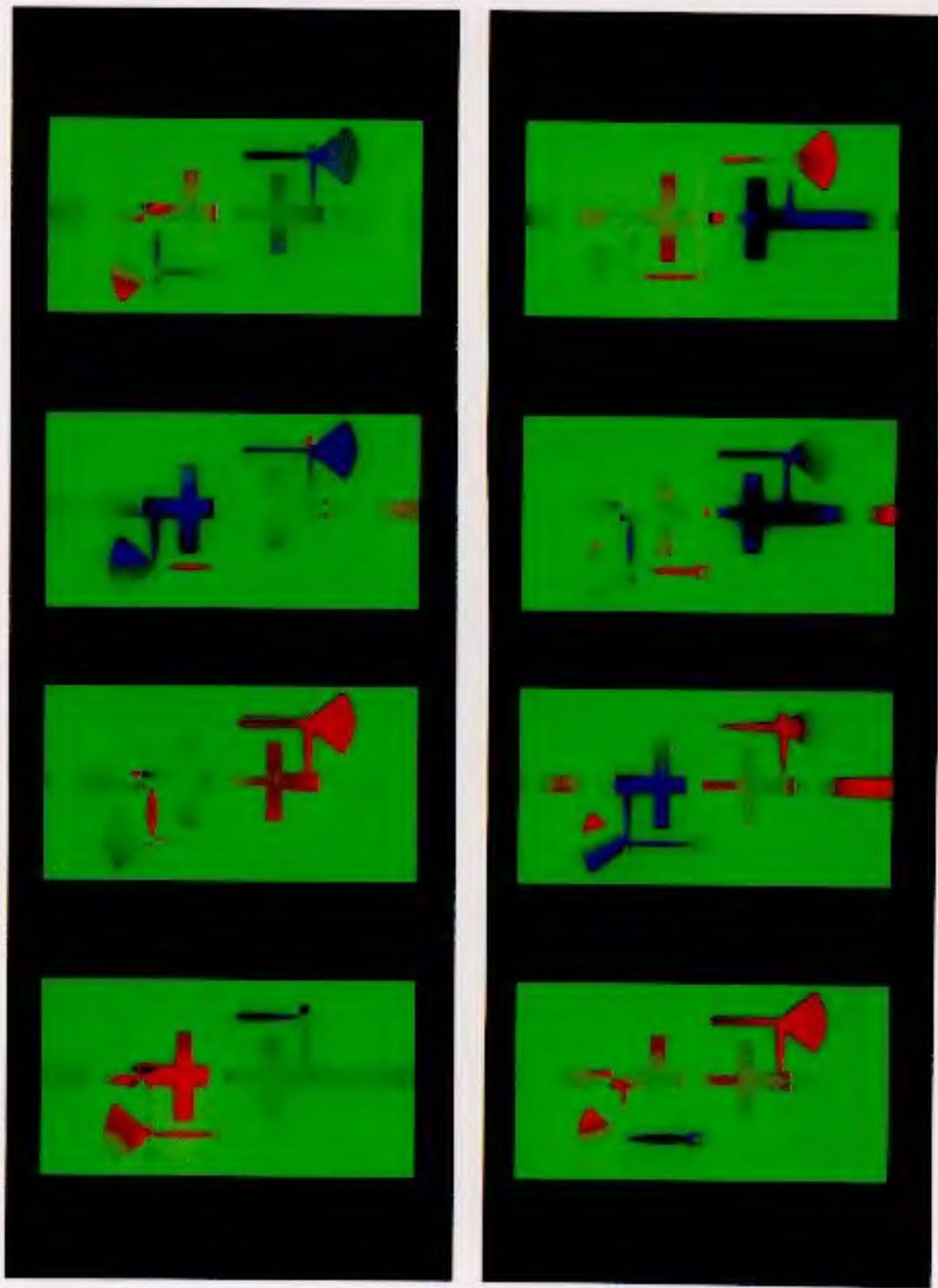


Fig. 1.4 Visualization of the FDTD-computed electromagnetic field within a 6-GHz microwave amplifier. The sequence of snapshots in time goes from top left to bottom right. (See also Figs. 15.16–15.18.)

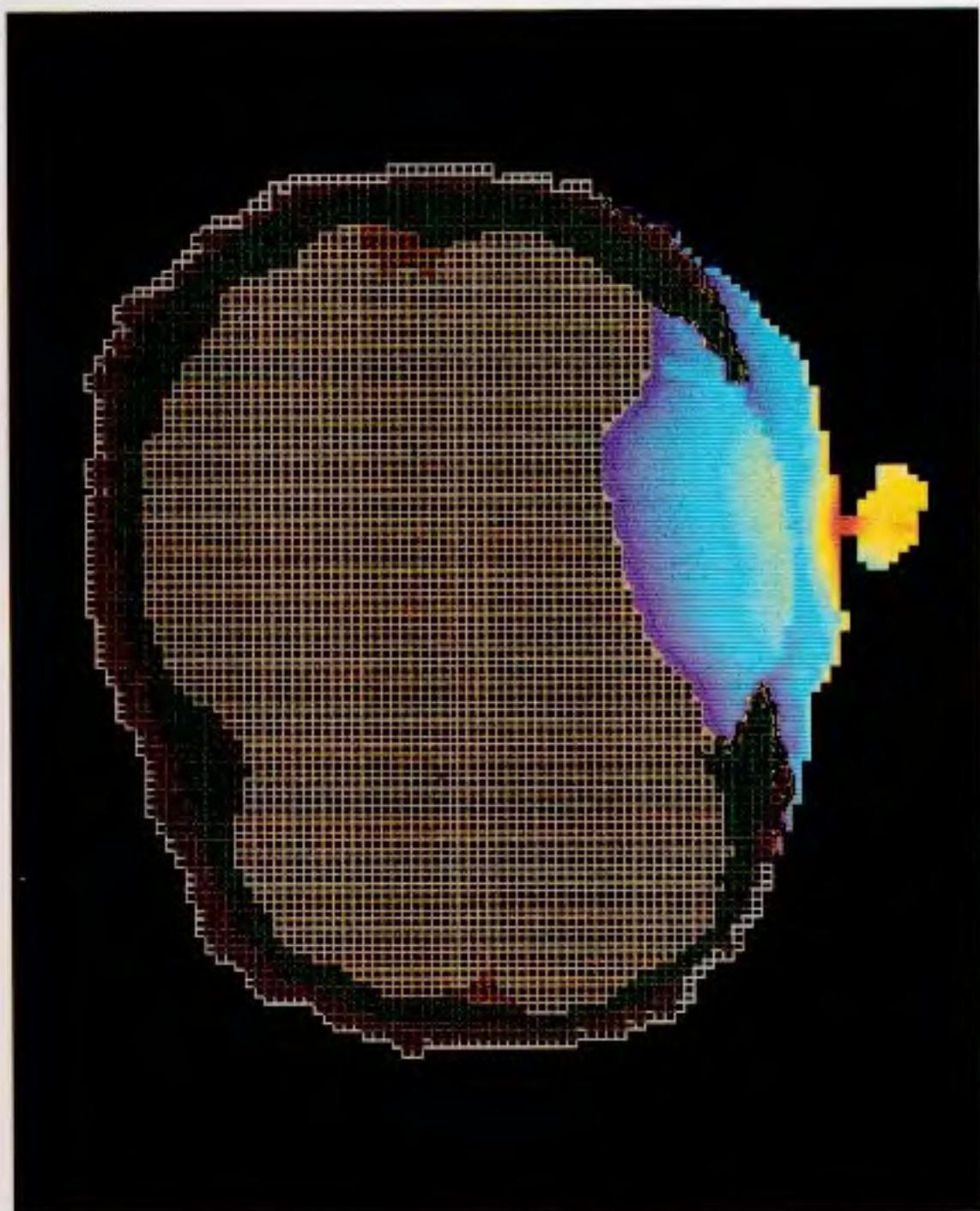


Fig. 1.5 Visualization of the FDTD-computed specific absorption rate (SAR) distribution in an ear-level plane for a 1,900-MHz cellular telephone held vertically against a tilted-head model. The antenna is a one-quarter wavelength whip. (See also Figs. 14.38 and 14.39.)

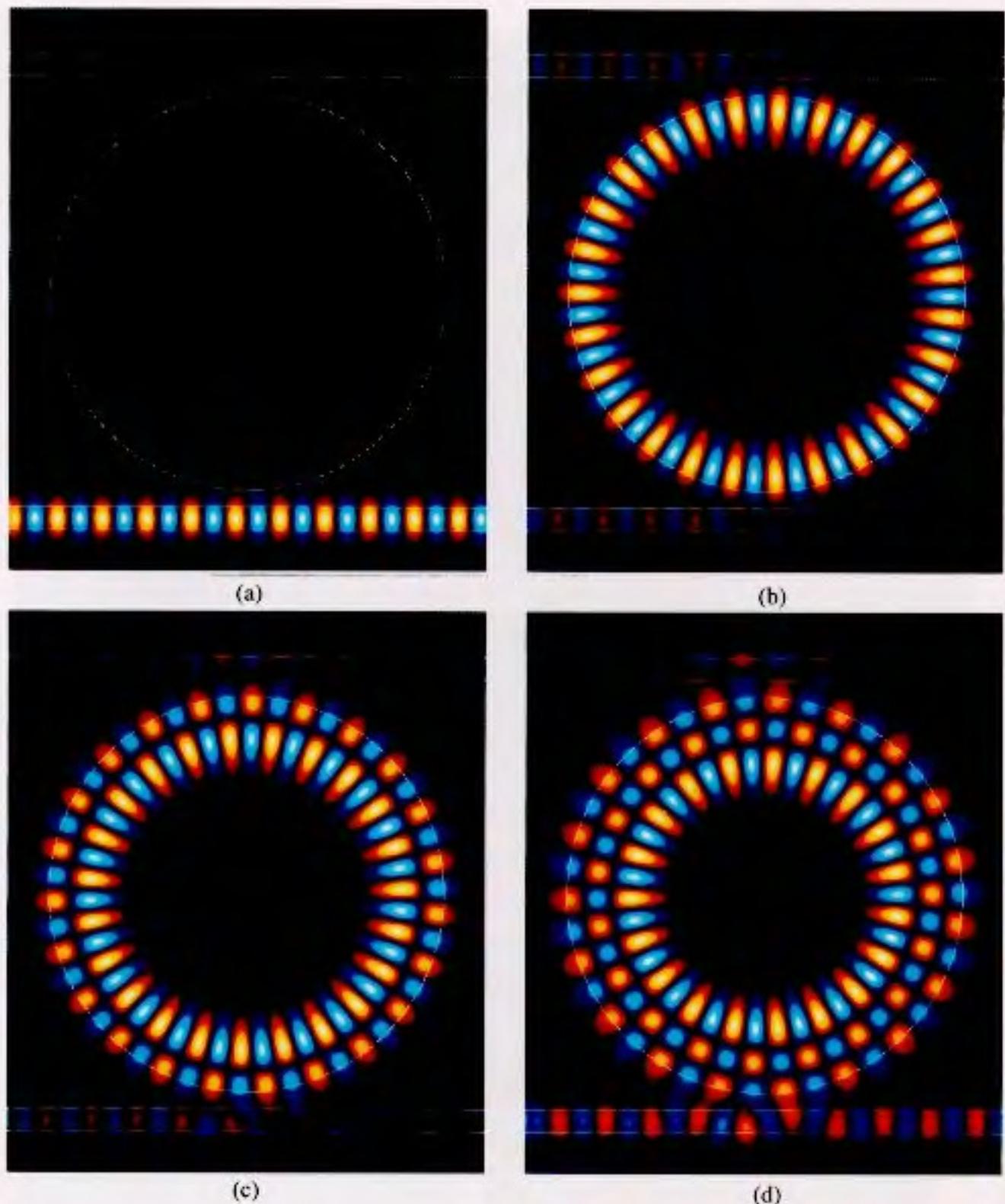


Fig. 1.6 Visualizations of the FDTD-computed sinusoidal steady-state electric field distributions in a 5.0- μm -diameter disk resonator coupled to straight 0.3- μm -wide optical waveguides for single-frequency excitations in the left side of the lower waveguide. (a) off-resonance signal; (b) on-resonance signal, first-order radial mode; (c) second-order radial-mode resonance; (d) third-order radial-mode resonance. (See also Figs. 16.21, 16.22, 16.32, and 16.34.)

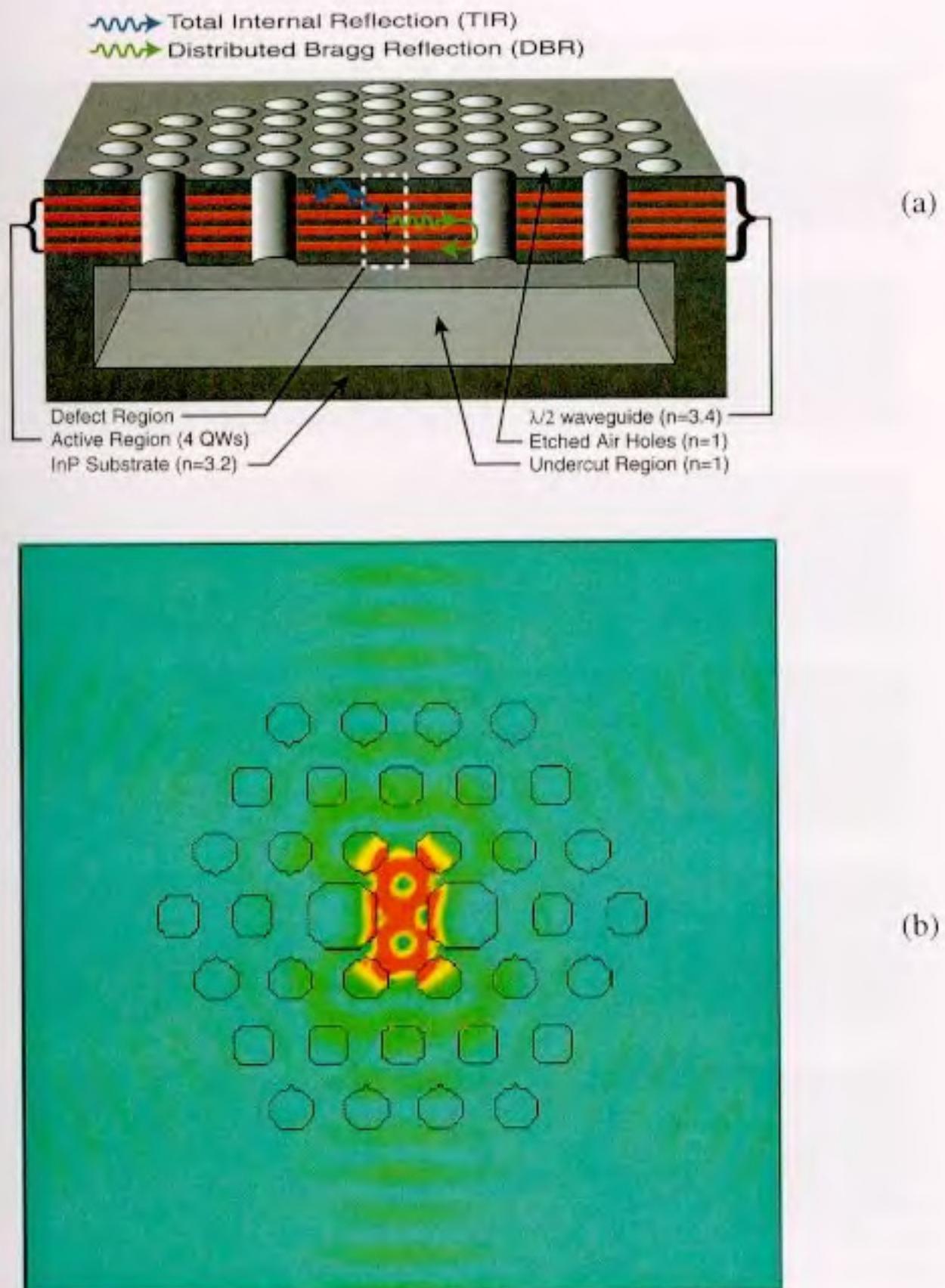


Fig. 1.7 Photonic bandgap microcavity laser: (a) geometry; (b) visualization of the FDTD-computed optical electric field along a planar cut through the middle of (a). *Source:* O. Painter et al., *Science*, June 11, 1999, pp. 1819-1821. (See also Figs. 16.18-16.20.)

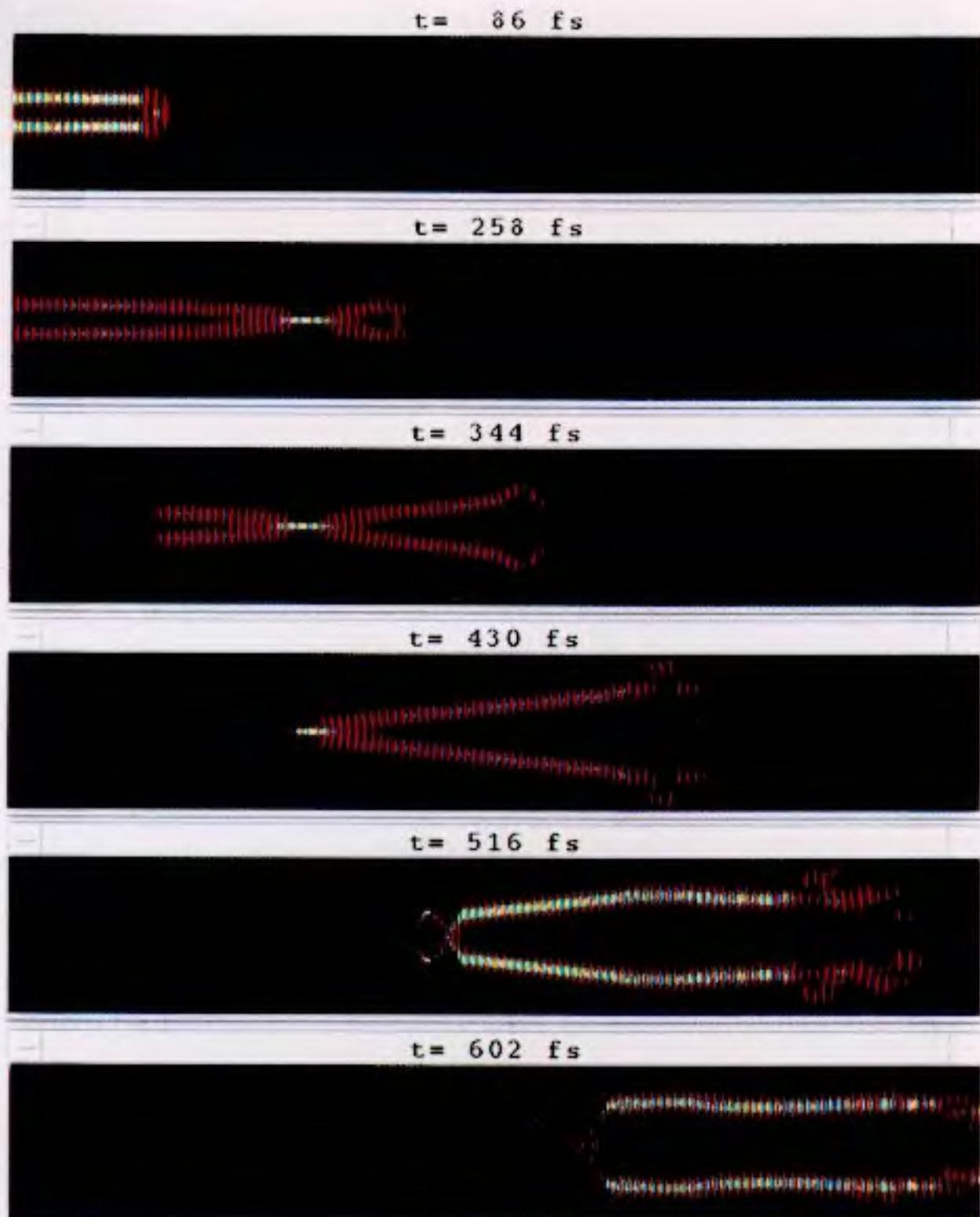


Fig. 1.8 Visualizations of the FDTD-computed electric field of equal-amplitude cophased 100-fs optical spatial solitons propagating in glass. This illustrates the dynamics of a potential all-optical “AND” gate, i.e., light switching light, that works on a time scale 1/10,000 that of existing electronic digital logic. (See also Fig. 9.5.)

1.6.2 Predictive Dynamic Range

For computational modeling of electromagnetic wave interaction structures using FDTD and related space-grid time-domain techniques, it is useful to consider the concept of predictive dynamic range. Let the power density of the primary (incident) wave in the space grid be P_0 W/m². Further, let the minimum observable power density of a secondary (scattered) wave be P_S W/m², where “minimum observable” means that the accuracy of the field computation degrades due to numerical artifacts to poorer than n dB (some desired figure of merit) at lower levels than P_S . Then, we can define the *predictive dynamic range* as $10 \log(P_0/P_S)$ dB.

This definition is well suited for FDTD and other space-grid time-domain codes for two reasons:

- It squares nicely with the concept of a “quiet zone” in an experimental anechoic chamber, which is intuitive to most electromagnetics engineers;
- It succinctly quantifies the fact that the desired numerical wave analogs propagating in the lattice exist in an additive noise environment due to nonphysical propagating wave analogs caused by the imperfect ABCs.

In addition to additive noise, the desired physical wave analogs undergo gradual progressive deterioration while propagating due to accumulating numerical dispersion artifacts, including phase velocity anisotropies and inhomogeneities within the mesh.

In the 1980s, researchers accumulated solid evidence for a predictive dynamic range on the order of 40–50 dB for FDTD codes. This value is reasonable if one considers the additive noise due to imperfect ABCs to be the primary limiting factor, since the analytical ABCs of this era provided effective outer-boundary reflection coefficients in the range of about 0.3%–3% (−30 to −50 dB).

The 1990s saw the emergence of powerful, entirely new classes of ABCs including Berenger’s PML [42–44]; the UPML of Sacks et al. and Gedney [46, 47]; and Ramahi’s COM and CCOM [51, 52]. These ABCs were shown to have effective outer-boundary reflection coefficients of better than −80 dB for impinging pulsed electromagnetic waves having ultrawideband spectra. Solid capabilities were demonstrated to terminate free-space lattices, multimoding and dispersive waveguiding structures, and lossy and dispersive materials.

However, for electrically large problems, the overall dynamic range may not reach the maximum permitted by these new ABCs because of inaccuracies due to accumulating numerical-dispersion artifacts generated by the basic grid-based solution of the curl equations. Fortunately, by the end of the 1990s, this problem was being attacked by a new generation of low-dispersion algorithms. Examples include the MRTD technique introduced by Krumpholz and Katehi [48] and the PSTD technique introduced by Liu [49, 50]. As a result of these advances, there is emerging the possibility of FDTD and related space-grid time-domain methods demonstrating predictive dynamic ranges of 80 dB or more in the first decade of the 21st century.

1.6.3 Scaling to Very Large Problem Sizes

Using FDTD and related methods, we can model electromagnetic wave interaction problems requiring the solution of considerably more than 10^8 field-vector unknowns. At this level of complexity, it is possible to develop detailed, three-dimensional models of complete engineering systems, including the following:

- Entire aircraft and missiles illuminated by radar at 1 GHz and above;
- Entire multilayer circuit boards and multichip modules for digital signal propagation, crosstalk, and radiation;
- Entire microwave and millimeter-wave amplifiers, including the active and passive circuit components and packaging;
- Entire integrated-optical structures, including lasers, waveguides, couplers, and resonators.

A key goal for such large models is to achieve algorithm / computer-architecture scaling such that for N field unknowns to be solved on M processors, we approach an order (N/M) scaling of the required computational resources.

We now consider the factors involved in determining the computational burden for the class of FDTD and related space-grid time-domain solvers.

1. *Number of volumetric grid cells, N .* The six vector electromagnetic field components located at each lattice cell must be updated at every time step. This yields by itself an order(N) scaling.
2. *Number of time steps, n_{\max} .* A self-consistent solution in the time domain mandates that the numerical wave analogs propagate over time scales sufficient to causally connect each portion of the structure of interest. Therefore, n_{\max} must increase as the maximum electrical size of the structure. In three dimensions, it can be argued that n_{\max} is a fractional power function of N such as $N^{1/3}$. Further, n_{\max} must be adequate to step through “ring-up” and “ring-down” times of energy storage features such as cavities. These features vary from problem to problem and cannot be ascribed a functional dependence relative to N .
3. *Cumulative propagation errors.* Additional computational burdens may arise due to the need for either progressive mesh refinement or progressively higher-accuracy algorithms to bound cumulative positional or phase errors for propagating numerical modes in progressively enlarged meshes. Any need for progressive mesh refinement would feed back to factor 1.

For most free-space problems, factors 2 and 3 are weaker functions of the size of the modeled structure than factor 1. This is because geometrical features at increasing electrical distances from each other become decoupled due to radiative losses by the electromagnetic waves propagating between these features. Further, it can be shown that replacing second-order accurate algorithms by higher-order versions sufficiently reduces numerical dispersion error to avoid the need for progressive mesh refinement for object sizes up to the order of 100 wavelengths. Overall, a computational burden of order($N \cdot n_{\max}$) = order($N^{4/3}$) is estimated for very large FDTD and related models.

1.7 EXAMPLES OF APPLICATIONS

This chapter continues with eight examples that illustrate both contemporary and emerging applications of FDTD computational electromagnetics modeling.

- *Radar-guided missile*
How do interactions between the missile seeker's horn antenna and radome generate errors in the angular location of a target?
- *High-speed computer circuit-board module*
Can digital circuits be upset at very-high clock rates due to parasitic coupling between signal paths?
- *High-speed computer multichip module (MCM)*
What is the inductance of a complex MCM power-distribution system?
- *Microwave amplifier*
What are the linear and nonlinear characteristics of a complete 6-GHz amplifier, including the effects of all circuit components and packaging?
- *Cellular telephone*
Can cellphones be designed to meet government safety standards for microwave exposure while meeting all other design goals?
- *Optical microdisk resonator*
How can proposed photonic integrated circuits best use these elements for ultrafast all-optical switching of signals based upon their wavelength?
- *Photonic bandgap (PBG) microcavity laser*
Can the world's smallest laser sources be based on PBG structures?
- *Colliding spatial solitons*
Can the world's fastest all-optical switches be built in ordinary glass?

Color visualizations are provided to help illustrate the nature of the problem-specific electrodynamics of these applications. Note that the discussion will refer to geometrical sketches and computer results contained in later chapters where each of these applications is presented in the context of specific FDTD techniques.

1.7.1 Radar-Guided Missile

We first consider an example of how the capability of the FDTD method to model structures having great volumetric complexity is useful for military defense applications. This example involves a long-standing defense technology problem that can cause unreliable homing of radar-guided missiles: the complex electromagnetic wave interactions between an antenna and its protective radome. These interactions generate errors in the perceived angular location of a target, thereby degrading the ability of the missile to guide itself to that target.

While the materials used in the construction of a missile radome are chosen to be as electromagnetically transparent as possible, the shape of the radome is usually dictated by aerodynamic considerations. Traditionally, the radome and its internal antenna are designed separately and their interaction is ignored. However, the current trend toward miniaturization and high-precision guidance is making this approach less valid. The FDTD method is well suited for use in analyzing both the antenna and its surrounding radome in the same computational model, thereby capturing the electromagnetic wave physics of the antenna-radome interaction.

Referring to the geometry sketches in Figs. 14.13 and 14.37(a) of Chapter 14, we consider an X-band pyramidal horn antenna having the characteristic dimensions $a = 7.30$ cm, $b = 5.32$ cm, and $D = 8$ cm. The missile is modeled as a 17.8-cm diameter, hollow, circular, perfect electric conductor (PEC) tube of wall thickness 7.5 mm. Its 33.0-cm-long nose section (the radome) is an ogive body of revolution formed from a lossless dielectric of permittivity $\epsilon_r = 4$. The nose cone is capped with a 2-cm long metal tip of thickness 3.75 mm.

We consider the case of illumination of the missile by an incident plane-wave pulse consisting of 6 sinusoidal cycles of a 10-GHz carrier. For this case, one plane of symmetry exists. The FDTD model uses a uniform space lattice composed of $449 \times 150 \times 300$ cubic Yee cells, each spanning 1.27 mm ($\approx \lambda_0/24$ at 10 GHz). Thus, the FDTD model solves for over 120-million field-vector components.

Fig. 1.1 is a color visualization of the E -field within and near the radome at two instants in time for the case of an incident plane wave propagating from right to left at 15° from boresight. In Fig. 1.1(a), the plane wave has entered the radome. A quasi-spherical, radially propagating, scattered field is observed due to the action of the metal tip and the radome surface. The propagation delay due to the reduction of the speed of light c inside the radome's dielectric wall is evident. In Fig. 1.1(b), the incident wave has encountered the horn antenna. A portion of the energy incident upon the horn antenna is scattered (structural-mode scattering). The waveguide attached to the horn is terminated in a matched load, and hence the antenna-mode scattering is zero. We note the packet of energy trapped inside the dielectric wall of the radome. The strong evanescent fields near the metal tip and the curvature of the radome allow the excitation of this trapped, guided wave. It is apparent that the visualizations resulting from the detailed FDTD model provide valuable engineering insights into the complex interactions of the missile body, radome, and internal horn antenna.

1.7.2 High-Speed Computer Circuit-Board Module

We next consider an example of how the capability of the FDTD method to model structures having great volumetric complexity is useful in the design of high-speed digital circuits. This example involves a critical problem that can cause unreliable operation: parasitic coupling of electronic pulses between signal paths. Most existing computer-aided circuit-design tools (primarily SPICE) are inadequate when digital clock speeds exceed about 500 MHz. These tools cannot deal with the physics of UHF/microwave electromagnetic wave energy transport along metal surfaces like ground planes, or in the air away from metal paths, which predominate above this frequency. Effectively, electronic digital systems develop substantial analog wave effects when clock rates are high enough, and full-vector Maxwell's equations solvers such as FDTD become necessary for understanding them.

In one of the most complex three-dimensional models of these effects so far, FDTD is used to simulate digital pulse propagation and crosstalk in a computer module consisting of a stack of four 22-layer circuit boards linked by three 50-pin connectors. The entire module is modeled with a uniform resolution of 0.004 inch, permitting each layer, via, and pin of each circuit board and connector to be modeled. 60-million field-vector components are solved, a factor of about 100 times larger than the capacity of the largest available SPICE or finite-element tool.

We consider the case of excitation of a single via pin in the top circuit board of the module by a subnanosecond pulse. Fig. 1.2(a) is a color visualization that shows strong early-time coupling of the H -field from the excited via pin to the adjacent unexcited pins, as seen in a vertical cut through the top board and connector of the stack. This represents undesired coupling to the digital circuits using these vias, which is capable of jamming the operation of these circuits at the leading edge of each logic pulse. Color video animations of the FDTD modeling results show an outwardly propagating electromagnetic wave within the top circuit board generated by the passage of the pulse down the via pin. Although the relatively intense H -field adjacent to the excited via is quite localized, moderate-level H -fields emanate throughout the entire transverse cross section of the board and link all of the adjacent via pins. In fact, the video shows repeated bursts of outward-propagating waves linking all points within the board as the pulse passes vertically through its multiple layers.

Fig. 1.2(b) is a color visualization of the magnitude and direction of late-time currents flowing along the vertical cross-section of the complete module for the excitation conditions of Fig. 1.2(a). These currents are calculated in a postprocessing step by numerically evaluating the curl of the H -field obtained from the FDTD model. Red denotes downward-directed current, while green denotes upward-directed current. At the time of this visualization, current has proceeded down the excited via through all four boards and all three connectors. However, upward-directed current is seen to flow on the adjacent vias. This represents an undesired ground-loop coupling to the digital circuits using these vias, an independent second coupling mode between digital signal paths that is capable of jamming the operation of these circuits.

1.7.3 Power-Distribution System for a High-Speed Computer Multichip Module

Our third example shows how the capability of the FDTD method to model structures having great volumetric complexity is useful in the design of power-distribution systems feeding a high-speed computer multichip module (MCM). Determining the equivalent inductance of a power-distribution system is of critical importance to successful package-level and system-level high-speed digital design. Physical and electrical design tradeoffs can be understood and optimal performance achieved if this inductance is known before commitment to hardware.

Low-inductance power-distribution systems usually have highly interdigitated geometries with globally distributed power and return paths. The inductance of these three-dimensional structures results from a complex interaction of the electromagnetic fields, and in general cannot be easily or accurately obtained using conventional circuit-design tools such as SPICE. Because the full-wave solution of Maxwell's equations by FDTD accounts for the physics of the structure's internal and external geometry to the resolution limit of the space lattice, a suitable analysis of the FDTD field calculations can be used to systematically calculate the inductance of arbitrarily shaped, three-dimensional structures. In addition, the full-wave FDTD solver frees the inductance calculation from any potentially limiting transverse electromagnetic wave (TEM) assumptions or approximations.

Referring to the geometry sketch of Fig. 15.4 of Chapter 15, we consider one such structure, a complex meshed printed-circuit-board (PCB) system having three power planes with current sources at the edges, five ground planes, and nine interdigitated vias connecting the power and ground planes. This model is based on a portion of an actual power-distribution structure that lies under a 2499-pin, 30-mm MCM. The MCM has a pad pitch of 0.5 mm and is built with copper/polyimide technology. The modeled portion covers a planar area of 3×3 mm and a via height of 1.02 mm. When the planes are meshed due to the signal-pin antipads, approximately 50% of the metal is removed. The meshed planes force the current to diverge around the antipad holes. Consequently, magnetic flux develops through the holes and the total inductance of the system is increased.

Fig. 1.3 is a color visualization of the FDTD-computed current density for this eight-layer power distribution system. This cutaway view shows the meshed ground plane wherein the MCM is located within the dotted rectangle. Using the simple postprocessing analysis discussed in Chapter 15, the inductance of the composite power-distribution system feeding the MCM is calculated to be approximately 10 pH.

This approach appears to be useful for general optimization of the physical configuration of a high-speed digital system to permit conformance with inductance ground rules. For example, in addition to optimizing the geometry of the meshed planes, this method can optimize the height of the vias that transport the signal from the power and ground planes of the PCB into the MCM.

1.7.4 Microwave Amplifier

It is an important and nontrivial conceptual leap to go from modeling only passive device packaging and interconnects, as discussed in Sections 1.4.2 and 1.4.3 above, to including in a self-consistent manner nonlinear active devices such as diodes, transistors, and logic gates in the electromagnetics model. The key to this process is linking circuit theory with electromagnetic field theory. In undergraduate electrical engineering programs, professors often tell students that circuit theory is a subset of field theory, but then promptly drop the connection because it is more convenient to proceed with analyzing the lumped devices of circuits rather than the distributed devices of fields. However, the circuits / fields connection is not simply of academic interest, and may be readily implemented for substantial practical gain.

To explore this connection, it is useful to remember just how circuit quantities (V , I , and Z) relate to field quantities (E - and H -fields). As discussed in Chapter 15, Faraday's and Ampere's Laws implemented in integral form within the FDTD space lattice are a natural means to accomplish this goal. Tests have shown that FDTD-computed circuit quantities calculated using appropriate contour integrals typically agree with textbook values on the order of 1% or better. Further, Norton's and Thevenin's equivalent circuits have been obtained for the FDTD lattice unit cell to enable a self-consistent connection of the FDTD-computed circuit quantities with circuit-analysis software such as SPICE. This results in SPICE subgrid models of transistors, logic gates, and passive components within the FDTD space lattice. In effect, *anything* that SPICE can model can be coupled into the FDTD lattice.

Referring to Figs. 15.16–15.18 of Chapter 15, we consider one such structure, a 6-GHz microwave amplifier wherein the nonlinear circuit characteristics of the metal-semiconductor (MESFET) transistor are incorporated into the three-dimensional FDTD model. Fig. 1.4 is a color visualization of snapshots of the FDTD-computed electromagnetic field within the amplifier. From the accompanying discussion in Section 15.12 of Chapter 15, it is apparent that the combined FDTD and nonlinear MESFET circuit model of the amplifier achieves very good results for small-signal S parameters, large-signal harmonic and intermodulation performance, and the effects of the amplifier's packaging upon its electrical stability relative to self oscillation.

We expect that continued progress in FDTD/SPICE will provide a novel and useful simulation tool for electrical engineers to obtain time-domain simulations of active, nonlinear, analog and digital circuits coupled directly to Maxwell's equations in three dimensions. This tool will be optimally applied when the speed of an active circuit is so high and its physical embedding is so compact and complex that the electromagnetic wave coupling, radiation, and ground-current artifacts of the embedding are crucial in its operation, and modeling the precise physical detail is required to properly understand the artifacts. In addition, FDTD/SPICE should have excellent capabilities in modeling circuit upset due to external natural and man-made electromagnetic insults such as lightning, EMP, and HPM.

1.7.5 Cellular Telephone

The 1990s have seen an explosive growth in the number and variety of personal wireless communications (PWC) devices available to the consumer. With market forces pushing for miniaturization and low cost, all aspects of cellular phones have been reengineered. A key requirement is to meet the government safety standard for microwave exposure to the user. Namely, the peak specific absorption rate (SAR) for any 1-g of tissue must be less than 1.6 W/kg.

FDTD simulations have become a powerful tool to meet the needs of engineers tasked with designing safe, efficient, compact antennas for PWC devices. This is because FDTD allows straightforward, accurate modeling of antenna near and far fields for essentially arbitrary configurations of inhomogeneous media. Such capability is mandatory for the success of any predictive tool aimed at addressing antenna interactions with the complex tissue structure of the human body.

To exploit the capability of FDTD to model extremely inhomogeneous dielectric materials, the head can be represented by a detailed anatomical model based on magnetic resonance imaging (MRI) of a human volunteer. A typical FDTD space lattice resolution in this case is $1.974 \times 1.974 \times 3.0$ mm. To avoid a staircased model of the cellphone handset and the antenna, it is desirable to rotate the head model instead of the handset through the desired angle at which the telephone is to be used.

Referring to Fig. 14.38 of Chapter 14, we consider one such MRI-based head model which is tilted forward by 30° . Here, a handset model placed vertically against the ear is tantamount to holding the telephone at an angle of 30° relative to a vertical (erect) head. Fig. 1.5 is a color visualization of the corresponding FDTD-computed SAR distribution in an ear-level plane. For this model, the telephone radiates 125 mW power at 1,900 MHz from a one-quarter-wavelength antenna. We see that the SAR is quite localized, with almost all of the absorbed energy occurring within the ear and the side of the head adjacent to the handset.

Experiments summarized in Table 14.2 of Chapter 14 compare peak 1-g SAR values obtained using the FDTD method with measured data for 10 cellphones, including five operating at 835 MHz and five at 1,900 MHz. The telephones include research test samples from several manufacturers and use a variety of antennas. The agreement between the measured and FDTD-calculated 1-g SARs is generally within $\pm 25\%$. This is noteworthy given the complexity of the inhomogeneous experimental and numerical models. Similarly, results shown in Fig. 14.40 of Chapter 14 indicate good agreement between FDTD-calculated and measured radiation patterns, with and without the human head adjacent to the cellphone.

We expect continued progress in FDTD modeling of PWC devices of all types, including (in addition to cellphones) laptop computers, personal digital assistants, Internet browsers, and satellite phones. Here, the potential for FDTD solutions of Maxwell's equations impacting everyday life is very large, since the number of people using PWC devices is climbing into the hundreds of millions worldwide, and electromagnetic waves are the means to link these devices to electronic networks.

1.7.6 Optical Microdisk Resonator

Waveguide-coupled microcavity ring and disk resonators are proposed components for filtering, routing, switching, modulation, and multiplexing/demultiplexing tasks in an emerging generation of photonic integrated circuits. Fig. 16.21 of Chapter 16 shows scanning electron microscope images of typical 10.5-micron diameter AlGaAs/GaAs microcavity ring and disk resonators coupled to 0.5-micron wide waveguides.

The quality of a resonator for optical communications is measured by its:

- Free-spectral range (FSR), or λ spacing between adjacent resonances;
- Finesse, or ratio of the FSR to the width of the resonance;
- Extinction ratio, or the ratio of the transmission at resonance to the off-resonance transmission.

The ideal resonator for a wavelength-division multiplexing (WDM) system has a wide FSR and a high finesse to accommodate many channels, high on-resonance transmission to minimize insertion loss, and a large extinction ratio to minimize crosstalk. A resonator having $\text{FSR} > 30 \text{ nm}$ is desirable for accommodating the WDM channels within the erbium-amplifier communications window. To achieve this goal, the diameter of the ring or disk must be just a few microns. In fact, semiconductor ring and disk resonators based on high-index contrast, strongly guiding optical waveguides can be designed with diameters as small as 1–2 microns with negligible bending loss. Current nanofabrication techniques permit the realization of such resonators coupled to submicron-width waveguides across submicron air gaps via evanescent waves.

Referring to Fig. 16.21, in the vertical direction each optical waveguiding structure consists of a 0.45-micron thick GaAs guiding layer with AlGaAs cladding layers above and below. Strong lateral confinement of the optical signal is achieved in the semiconductor waveguide by etching air gaps down through the lower cladding layer. The structure is weakly guiding in the vertical direction to allow for the input coupling of light from an external fiber or planar-emitting device.

The etching process results in a 10–20 nm surface roughness of the sidewalls of the optical waveguides, which scatters light propagating within. To directly model this important fabrication artifact, a grid cell of comparable size is used in the FDTD simulation. The fine grid-cell size also permits detailed modeling of the evanescent electromagnetic wave coupling across the air gaps between the coupling waveguides and the resonators, even down to the minimum distance that can be reliably fabricated with current technology, on the order of 100 nm.

Fig. 16.22 of Chapter 16 shows a typical FDTD modeling geometry for a microcavity ring or disk resonator. Optical waveguides WG1 and WG2 serve as evanescent wave input and output couplers. In the FDTD model, a fundamental mode is sourced at the left end of WG1. Field components in the simulated waveguide mode are recorded as it propagates past cross-sections P1 and P2 in WG1, and P3 in the

microring or microdisk. By taking discrete Fourier transforms of the time waveforms of the fields, calculating the Poynting power densities along the cross-sections, and summing, the longitudinal power flux passing through each cross-section is obtained as a function of frequency. In this manner, a single FDTD computer run provides a complete spectral characterization of the coupling efficiency (the flux through P3 normalized by the incident flux through P1) and the transmittance (the normalized flux through P2). Where benchmark data are available, the accuracy of this modeling approach is found to be excellent. For example, power balances in the propagating mode computed in this manner are found to hold within parts per ten thousand.

In the present example, we consider a 5.0-micron diameter disk resonator with a 0.232-micron air gap between the disk and each adjacent straight 0.3-micron wide waveguide. To observe the whispering-gallery modes of the microdisk resonator, single-frequency FDTD simulations are performed at the resonance frequencies of specific modes of interest.

Fig. 1.6(a) is a color visualization of the FDTD-computed sinusoidal steady-state E -field distribution for excitation at a nonresonant frequency, 193.4 THz ($\lambda = 1.55$ microns). Calculations made from the field data for this case show that 99.98% of the power in the incident signal remains in WG1.

Fig. 1.6(b) is the corresponding visualization at the resonant frequency of a first-order radial whispering-gallery mode, 189.2 THz ($\lambda = 1.585$ microns). Here, the field pattern within the microdisk shows the effect of resonant field enhancement. Since the color scale here is normalized to the peak resonant fields, the fields in WG1 and WG2 appear to be weak. In fact, they are at the same level as the fields in WG1 in Fig. 1.6(a). For this on-resonance case, 99.79% of the incident power in WG1 switches to WG2. A comparison between Figs. 1.6(a) and 1.6(b) indicates the potential for high extinction ratios and low crosstalk between channels when using the microdisk device as a passive, wavelength-selective switch.

Figs. 1.6(c) and 1.6(d) are the corresponding visualizations at, respectively, the resonant frequency of a second-order radial whispering-gallery mode, 191.3 THz ($\lambda = 1.567$ microns), and the resonant frequency of a third-order radial whispering-gallery mode, 187.8 THz ($\lambda = 1.596$ microns). These visualizations show how each resonant whispering-gallery mode is confined to an annulus around the perimeter of the disk wherein the width of the annulus increases as the radial order increases. A key goal of current design efforts is to suppress such higher order modes of the microdisks to allow their effective usage as passive WDM devices having low crosstalk across a wide spectrum, or as active single-mode laser sources.

In the near future, we expect continued progress in FDTD modeling of optical waveguides and resonators of all types. The engineering design of proposed photonic circuits will benefit from full-wave FDTD modeling in the same way as microwave circuit design. This is because both microwave circuits and proposed photonic circuits operate in approximately the same regime of the size of their components relative to the wavelength of the microwave or optical signal being processed. Further, both use linear passive and nonlinear active components capable of direct FDTD simulation.

1.7.7 Photonic Bandgap Microcavity Laser

Photonic crystals are artificial structures that have a periodic variation of the refractive index in one, two, or three dimensions. Analogous to the energy gap in pure semiconductor crystals in which electrons are forbidden, these photonic bandgap (PBG) structures have a frequency stopband over which there is no transmission of electromagnetic waves. Similar to a donor or acceptor state in a doped semiconductor, a small defect introduced into the photonic crystal creates a resonant mode at a frequency that lies inside the bandgap. The defect in the periodic array behaves as a microcavity resonator.

As discussed in Section 16.6 of Chapter 16, FDTD modeling has been used to design the world's smallest laser sources based on PBG structures [57]. Here, light is confined to a single defect of a nanofabricated two-dimensional photonic crystal. The ability to fabricate compact lateral microcavities is important for the construction of high-density multiwavelength optical circuits. The advantage of using a photonic crystal is the inherent flexibility in geometry which allows fine tuning of the defect-mode radiation pattern as well as the emission wavelength. The compact size and high spontaneous-emission coupling factor of the defect microcavity also make it interesting as a low-noise, low-threshold light source. In addition, microcavities based upon nanofabricated photonic crystals may be useful where crystal growth of high-index contrast mirrors are limited, such as in long-wavelength vertical-cavity surface-emitting lasers or blue-green gallium nitride-based devices.

Fig. 1.7(a) illustrates how light is contained inside the microcavity. First, a $\lambda/2$ high-index slab is used to trap electromagnetic fields in the vertical direction by way of total-internal reflection (TIR) at the air-slab interface. Second, the light is localized in-plane using a fabricated two-dimensional photonic crystal consisting of a hexagonal array of air holes etched into the slab. The periodic variation in the refractive index gives rise to Bragg scattering, which generates forbidden energy gaps in the in-plane electromagnetic wave dispersion relation. Thus, the photonic crystal provides an energy barrier for the propagation of guided electromagnetic waves having frequencies that lie within the band gap. In the simplest structure, a single air hole is removed from the photonic crystal, thereby forming a resonant microcavity. The light energy in the resonant mode is highly spatially localized in the defect region, and can escape only by either tunneling through the two-dimensional photonic crystal or by impinging on the air-slab interface at a sufficiently large angle to leak out in the vertical direction.

The defect laser cavities are fabricated in indium gallium arsenic phosphide (InGaAsP) using metalorganic chemical vapor deposition. Here, the active region consists of four 9-nm quantum wells separated by 20-nm quaternary barriers with a 1.22-micron band gap. The quantum-well emission wavelength is designed for 1.55 microns at room temperature.

Fig. 16.19 of Chapter 16 is a scanning electron micrograph that shows a top view of the fabricated microcavity. The important parameters in determining the properties of this optical cavity are the refractive index n_{slab} of the dielectric slab, the thickness d

of the slab, the interhole spacing a , and the radius r of the air holes. A pair of degenerate dipole modes arises upon removing a single air hole in the slab: the x -dipole mode and the y -dipole mode. It is desirable to suppress the x -dipole mode to achieve single-mode operation. This is accomplished by enlarging the air holes to either x -side of the defect to the radius $r' > r$, which pushes the x -dipole mode frequency outside of the band gap of the photonic crystal. Further, this pair of enlarged air holes tunes the y -dipole mode frequency so as to maximize the cavity Q . The key dimensions of the laser cavity as measured from the micrograph of Fig. 16.19 are $a = 515$ nm, $r = 180$ nm, $r' = 240$ nm, and $d = 210$ nm. These are incorporated into the three-dimensional FDTD model.

Fig. 1.7(b) is a color visualization of the magnitude of the E -field calculated using FDTD along a planar cut through the middle of the InGaAsP slab. FDTD simulations of this cavity with $n_{\text{slab}} = 3.4$ give an in-plane band gap extending from 1.25–1.72 microns, a resonant wavelength of 1.509 microns, a quality factor Q of 250, and an effective modal volume of 0.03 cubic microns. Nearly all of the power is emitted vertically due to the band gap of the surrounding crystal. Experimental realization of this microcavity laser indicates a lasing wavelength of 1.504 microns, very close to the FDTD-predicted value. This is achieved at an operating temperature of 143°K using a microscope to focus a 6.75-mW (threshold) optical pump beam of 0.830-micron wavelength upon the cavity. Ongoing work involves optimizing this laser cavity by tailoring the ratios r/a and r'/a to elevate the Q , which theoretically can be increased above 1,500. This would significantly lower pump threshold values and permit the device to operate at or near room temperature.

In the near future, we expect FDTD modeling to become widely used in the design of lasers such as the one discussed above, incorporating even more details of the physics of light emission by such devices. With the capability of FDTD to accurately model active lasing devices as well as passive PBG structures, optical waveguides, couplers, and cavities as seen from this example and the previous example, it is almost certain that large-scale FDTD modeling will be used to design complete photonic integrated circuits.

1.7.8 Colliding Spatial Solitons

In electrical engineering, the phrase “dc to daylight” has been often used to describe electronic systems having the property of very wide bandwidth. Of course, no one actually meant that the system in question could produce or process signals over this frequency range. It just couldn’t be done. Or could it?

In fact, a simple Fourier analysis argument shows that recent optical systems that generate laser pulses down to 10 fs in duration approach this proverbial bandwidth. From a technology standpoint, it is clear that controlling or processing these short pulses involves understanding the nature of their interactions with materials over nearly “dc to daylight,” and very likely in high-beam-intensity regimes where material nonlinearity can play an important role. A key factor here is material dispersion,

having two components: (1) linear dispersion, the variation of the material's index of refraction with frequency at low optical power levels; and (2) nonlinear dispersion, the variation of the frequency-dependent refractive index with optical power.

During the 1990s, advances in FDTD techniques provided the basis for modeling both linear and nonlinear dispersions over ultrawide bandwidths. An example of the possibilities for such FDTD modeling is shown in Fig. 1.8, which is a color visualization of the dynamics of a potential femtosecond all-optical switch using colliding spatial solitons. From a pair of optical waveguides on the left side, this switch would inject cophased 100-fs signal and control solitons having a 0.65-micron beamwidth into glass, a Kerr-type nonlinear material. In the absence of the control soliton, the signal soliton would propagate to the right with zero deflection, and then be collected by a receiving waveguide. In the presence of the control soliton, the signal soliton would be laterally deflected to an alternate collecting waveguide after colliding with the control soliton. This would provide the action of an all-optical "AND" gate working on a time scale about 1/10,000 that of existing electronic digital logic.

The FDTD method achieves robustness for nonlinear optical problems by retaining the optical carrier and solving for fundamental quantities, the optical \vec{E} and \vec{H} fields in space and time. Note that most previous approaches in nonlinear optics calculated a nonphysical scalar carrier-envelope function after discarding much of the full-wave physics contained in Maxwell's equations. In addition, FDTD enforces the \vec{E} and \vec{H} boundary conditions at all material interfaces in the time scale of the carrier, whether or not the media are dispersive or nonlinear. As a result, it is almost completely general. As illustrated in Fig. 1.8, it may be possible to use FDTD to design all-optical switching circuits (light switching light) attaining speeds 10,000 times faster than those of the best semiconductor circuits today. The implications may be profound for the realization of "optronics," a proposed successor technology to electronics in the 21st century, that would integrate optical-fiber interconnects and optical microchips into systems of unimaginable information-processing capability.

1.8 CONCLUSIONS

Whereas large-scale solutions of Maxwell's equations have been motivated in the past primarily by the requirements of military defense, the entire field of computational electrodynamics is shifting rapidly toward important commercial applications in high-speed communications and computing that will touch everyone in their daily lives. Ultimately, this will favorably impact the economic well-being of nations as well as their military security.

In fact, the large-scale solution of Maxwell's equations for electromagnetic wave phenomena using FDTD and similar grid-based time-domain approaches may be fundamental to the advancement of electrical and computer engineering technology as we continue to push the envelope of the ultracomplex and the ultrafast. Maxwell's equations provide the physics of electromagnetic wave phenomena from dc to light,

and their accurate modeling is essential to understand high-speed signal effects having wave-transport behavior. A key goal is the computational unification of electromagnetic waves; charge transport in transistors, Josephson junctions, and electro-optic devices; surface and volumetric wave dispersions (including those of superconductors); and nonlinearities due to quantum effects. Then we can attack a broad spectrum of important problems to advance electrical and computer engineering and directly benefit our society.

REFERENCES

- [1] Keller, J. B., "Geometrical theory of diffraction," *J. Optical Society of America*, Vol. 52, 1962, pp. 116–130.
- [2] Kouyoumjian, R. G., and P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface," *Proc. IEEE*, Vol. 62, 1974, pp. 1448–1461.
- [3] Harrington, R. F., *Field Computation by Moment Methods*, New York: Macmillan, 1968.
- [4] Umashankar, K. R., "Numerical analysis of electromagnetic wave scattering and interaction based on frequency-domain integral equation and method of moments techniques," *Wave Motion*, Vol. 10, 1988, pp. 493–525.
- [5] Song, J., and W. C. Chew, "The Fast Illinois Solver Code: Requirements and scaling properties," *IEEE Computational Science & Engineering*, Vol. 5, July-Sept. 1998, pp. 19–23.
- [6] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas and Propagation*, Vol. 14, 1966, pp. 302–307.
- [7] Shlager, K. L., and J. B. Schneider, "A Survey of the Finite-Difference Time-Domain Literature," Chap. 1 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, ed., Norwood, MA: Artech House, 1998.
- [8] Taflove, A., and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory and Techniques*, Vol. 23, 1975, pp. 623–630.
- [9] Taflove, A., and M. E. Brodwin, "Computation of the electromagnetic fields and induced temperatures within a model of the microwave-irradiated human eye," *IEEE Trans. Microwave Theory and Techniques*, Vol. 23, 1975, pp. 888–896.
- [10] Holland, R., "Threde: a free-field EMP coupling and scattering code," *IEEE Trans. Nuclear Science*, Vol. 24, 1977, pp. 2416–2421.
- [11] Kunz, K. S., and K. M. Lee, "A three-dimensional finite-difference solution of the external response of an aircraft to a complex transient EM environment I: The method and its implementation," *IEEE Trans. Electromagnetic Compatibility*, Vol. 20, 1978, pp. 328–333.
- [12] Taflove, A., "Application of the finite-difference time-domain method to sinusoidal steady-state electromagnetic penetration problems," *IEEE Trans. Electromagnetic Compatibility*, Vol. 22, 1980, pp. 191–202.
- [13] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagnetic Compatibility*, Vol. 23, 1981, pp. 377–382.

- [14] Umashankar, K. R., and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Trans. Electromagnetic Compatibility*, Vol. 24, 1982, pp. 397–405.
- [15] Taflove, A., and K. R. Umashankar, "Radar cross section of general three-dimensional scatterers," *IEEE Trans. Electromagnetic Compatibility*, Vol. 25, 1983, pp. 433–440.
- [16] Choi, D. H., and W. J. Hoefer, "The finite-difference time-domain method and its application to eigenvalue problems," *IEEE Trans. Microwave Theory and Techniques*, Vol. 34, 1986, pp. 1464–1470.
- [17] Kriegsmann, G. A., A. Taflove, and K. R. Umashankar, "A new formulation of electromagnetic wave scattering using an on-surface radiation boundary condition approach," *IEEE Trans. Antennas and Propagation*, Vol. 35, 1987, pp. 153–161.
- [18] Moore, T. G., J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, "Theory and application of radiation boundary operators," *IEEE Trans. Antennas and Propagation*, Vol. 36, 1988, pp. 1797–1812.
- [19] Umashankar, K. R., A. Taflove, and B. Beker, "Calculation and experimental validation of induced currents on coupled wires in an arbitrary shaped cavity," *IEEE Trans. Antennas and Propagation*, Vol. 35, 1987, pp. 1248–1257.
- [20] Taflove, A., K. R. Umashankar, B. Beker, F. A. Harfoush, and K. S. Yee, "Detailed FDTD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens," *IEEE Trans. Antennas and Propagation*, Vol. 36, 1988, pp. 247–257.
- [21] Jurgens, T. G., A. Taflove, K. R. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas and Propagation*, Vol. 40, 1992, pp. 357–366.
- [22] Cangellaris, A. C., C.-C. Lin, and K. K. Mei, "Point-matched time-domain finite element methods for electromagnetic radiation and scattering," *IEEE Trans. Antennas and Propagation*, Vol. 35, 1987, pp. 1160–1173.
- [23] Shankar, V., A. H. Mohammadian, and W. F. Hall, "A time-domain finite-volume treatment for the Maxwell equations," *Electromagnetics*, Vol. 10, 1990, pp. 127–145.
- [24] Madsen, N. K., and R. W. Ziolkowski, "A three-dimensional modified finite volume technique for Maxwell's equations," *Electromagnetics*, Vol. 10, 1990, pp. 147–161.
- [25] Sullivan, D. M., O. P. Gandhi, and A. Taflove, "Use of the finite-difference time-domain method in calculating EM absorption in man models," *IEEE Trans. Biomedical Engineering*, Vol. 35, 1988, pp. 179–186.
- [26] Zhang, X., J. Fang, K. K. Mei, and Y. Liu, "Calculation of the dispersive characteristics of microstrips by the time-domain finite-difference method," *IEEE Trans. Microwave Theory and Techniques*, Vol. 36, 1988, pp. 263–267.
- [27] Kashiwa, T., and I. Fukai, "A treatment by FDTD method of dispersive characteristics associated with electronic polarization," *Microwave and Optics Technologies Letters*, Vol. 3, 1990, pp. 203–205.
- [28] Luebbers, R., F. Hunsberger, K. Kunz, R. Standler, and M. Schneider, "A frequency-dependent finite-difference time-domain formulation for dispersive materials," *IEEE Trans. Electromagnetic Compatibility*, Vol. 32, 1990, pp. 222–229.
- [29] Joseph, R. M., S. C. Hagness, and A. Taflove, "Direct time integration of Maxwell's equations in linear dispersive media with absorption for scattering and propagation of femtosecond electromagnetic pulses," *Optics Letters*, Vol. 16, 1991, pp. 1412–1414.

- [30] Maloney, J. G., G. S. Smith, and W. R. Scott, Jr., "Accurate computation of the radiation from simple antennas using the finite-difference time-domain method," *IEEE Trans. Antennas and Propagation*, Vol. 38, 1990, pp. 1059–1065.
- [31] Katz, D. S., A. Taflove, M. J. Piket-May, and K. R. Umashankar, "FDTD analysis of electromagnetic wave radiation from systems containing horn antennas," *IEEE Trans. Antennas and Propagation*, Vol. 39, 1991, pp. 1203–1212.
- [32] Tirkas, P. A., and C. A. Balanis, "Finite-difference time-domain technique for radiation by horn antennas," *1991 IEEE Antennas and Propagation Society International Symposium Digest*, Vol. 3, 1991, pp. 1750–1753.
- [33] Sano, E., and T. Shibata, "Fullwave analysis of picosecond photoconductive switches," *IEEE J. Quantum Electronics*, Vol. 26, 1990, pp. 372–377.
- [34] El-Ghazaly, S. M., R. P. Joshi, and R. O. Grondin, "Electromagnetic and transport considerations in subpicosecond photoconductive switch modeling," *IEEE Trans. Microwave Theory and Techniques*, Vol. 38, 1990, pp. 629–637.
- [35] Goorjian, P. M., and A. Taflove, "Direct time integration of Maxwell's equations in nonlinear dispersive media for propagation and scattering of femtosecond electromagnetic solitons," *Optics Letters*, Vol. 17, 1992, pp. 180–182.
- [36] Ziolkowski, R. W., and J. B. Judkins, "Full-wave vector Maxwell's equations modeling of self-focusing of ultra-short optical pulses in a nonlinear Kerr medium exhibiting a finite response time," *J. Optical Society of America B*, Vol. 10, 1993, pp. 186–198.
- [37] Joseph, R. M., P. M. Goorjian, and A. Taflove, "Direct time integration of Maxwell's equations in 2-D dielectric waveguides for propagation and scattering of femtosecond electromagnetic solitons," *Optics Letters*, Vol. 18, 1993, pp. 491–493.
- [38] Joseph, R. M., and A. Taflove, "Spatial soliton deflection mechanism indicated by FDTD Maxwell's equations modeling," *IEEE Photonics Technology Letters*, Vol. 2, 1994, pp. 1251–1254.
- [39] Sui, W., D. A. Christensen, and C. H. Durney, "Extending the two-dimensional FDTD method to hybrid electromagnetic systems with active and passive lumped elements," *IEEE Trans. Microwave Theory and Techniques*, Vol. 40, 1992, pp. 724–730.
- [40] Toland, B., B. Houshmand, and T. Itoh, "Modeling of nonlinear active regions with the FDTD method," *IEEE Microwave and Guided Wave Letters*, Vol. 3, 1993, pp. 333–335.
- [41] Thomas, V. A., M. E. Jones, M. J. Piket-May, A. Taflove, and E. Harrigan, "The use of SPICE lumped circuits as sub-grid models for FDTD high-speed electronic circuit design," *IEEE Microwave and Guided Wave Letters*, Vol. 4, 1994, pp. 141–143.
- [42] Berenger, J. P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics*, Vol. 114, 1994, pp. 185–200.
- [43] Katz, D. S., E. T. Thiele, and A. Taflove, "Validation and extension to three dimensions of the Berenger PML absorbing boundary condition for FDTD meshes," *IEEE Microwave and Guided Wave Letters*, Vol. 4, 1994, pp. 268–270.
- [44] Reuter, C. E., R. M. Joseph, E. T. Thiele, D. S. Katz, and A. Taflove, "Ultrawideband absorbing boundary condition for termination of waveguiding structures in FDTD simulations," *IEEE Microwave and Guided Wave Letters*, Vol. 4, 1994, pp. 344–346.
- [45] Gedney, S., and F. Lansing, "A parallel planar generalized Yee algorithm for the analysis of microwave circuit devices," *Int. J. for Numerical Modeling (Electronic Networks, Devices, and Fields)*, Vol. 8, May-Aug. 1995, pp. 249–264.

- [46] Sacks, Z. S., D. M. Kingsland, R. Lee, and J. F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Trans. Antennas and Propagation*, Vol. 43, 1995, pp. 1460–1463.
- [47] Gedney, S. D., "An anisotropic perfectly matched layer absorbing media for the truncation of FDTD lattices," *IEEE Trans. Antennas and Propagation*, Vol. 44, 1996, pp. 1630–1639.
- [48] Krumpholz, M., and L. P. B. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory and Techniques*, Vol. 44, 1996, pp. 555–572.
- [49] Liu, Q. H., *The PSTD Algorithm: A Time-Domain Method Requiring Only Two Grids Per Wavelength*, New Mexico State Univ., Las Cruces, NM, Technical Report NMSU-ECE96-013, Dec. 1996.
- [50] Liu, Q. H., "The pseudospectral time-domain (PSTD) method: A new algorithm for solutions of Maxwell's equations," *Proc. 1997 IEEE Antennas and Propagation Society Intl. Symposium*, (IEEE catalog no. 97CH36122), July 13–18, 1997, Montreal, Canada, Vol. 1, pp. 122–125.
- [51] Ramahi, O. M., "The complementary operators method in FDTD simulations," *IEEE Antennas and Propagation Magazine*, Vol. 39, no. 6, Dec. 1997, pp. 33–45.
- [52] Ramahi, O. M., "The concurrent complementary operators method for FDTD mesh truncation," *IEEE Trans. Antennas and Propagation*, Vol. 46, 1998, pp. 1475–1482.
- [53] Dey, S., and R. Mittra, "A locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *IEEE Microwave and Guided Wave Letters*, Vol. 7, 1997, pp. 273–275.
- [54] Dey, S., and R. Mittra, "A modified locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *Microwave and Optical Technology Letters*, Vol. 17, 1998, pp. 349–352.
- [55] Maloney, J. G., and M. P. Kesler, "Analysis of Periodic Structures," Chapter 6 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, ed., Norwood, MA: Artech House, 1998.
- [56] Schneider, J. B., and C. L. Wagner, "FDTD dispersion revisited: Faster-than-light propagation," *IEEE Microwave and Guided Wave Letters*, Vol. 9, 1999, pp. 54–56.
- [57] Painter, O., R. K. Lee, A. Scherer, A. Yariv, J. D. O'Brien, P. D. Dapkus, and I. Kim, "Two-dimensional photonic band-gap defect mode laser," *Science*, Vol. 284, 11 June 1999, pp. 1819–1821.
- [58] Zheng, F., Z. Chen, and J. Zhang, "Towards the development of a three-dimensional unconditionally stable finite-difference time-domain method," *IEEE Trans. Microwave Theory and Techniques*, Vol. 48, 2000 (in press).
- [59] Zheng, F., and Z. Chen, "Numerical dispersion analysis of the unconditionally stable 3D ADI-FDTD method," *IEEE Trans. Microwave Theory and Techniques*, Vol. 48, 2000 (in press).

Chapter 2

The One-Dimensional Scalar Wave Equation

2.1 INTRODUCTION

In this chapter we consider the numerical FDTD solution of the most basic partial differential equation that describes wave motion, the one-dimensional scalar wave equation. The analytical propagating-wave solutions are first obtained. Then finite differences are introduced and applied to the wave equation, leading to introductory discussions of numerical dispersion, numerical phase velocity, the “magic” time step, and numerical stability. The discussions of this chapter serve as the basis for later work with similar concepts regarding FDTD analysis of the vector Maxwell’s equations in two and three dimensions.

2.2 PROPAGATING-WAVE SOLUTIONS

Consider the one-dimensional scalar wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (2.1)$$

where $u = u(x, t)$. What are the possible solutions? Consider functions of the type

$$u(x, t) = F(x + ct) + G(x - ct) \quad (2.2)$$

where F and G are arbitrary. Let’s partially differentiate these twice with respect to t and x :

$$\begin{aligned}\frac{\partial u}{\partial t} &= \underbrace{\frac{dF(x+ct)}{d(x+ct)}}_{F'} \cdot \underbrace{\frac{\partial(x+ct)}{\partial t}}_c + \underbrace{\frac{dG(x-ct)}{d(x-ct)}}_{G'} \cdot \underbrace{\frac{\partial(x-ct)}{\partial t}}_{-c} \\ &= cF'(x+ct) - cG'(x-ct)\end{aligned}\quad (2.3a)$$

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} &= c^2 F''(x+ct) - c(-c) G''(x-ct) \\ &= c^2 F''(x+ct) + c^2 G''(x-ct)\end{aligned}\quad (2.3b)$$

$$\begin{aligned}\frac{\partial u}{\partial x} &= \underbrace{\frac{dF(x+ct)}{d(x+ct)}}_{F'} \cdot \underbrace{\frac{\partial(x+ct)}{\partial x}}_1 + \underbrace{\frac{dG(x-ct)}{d(x-ct)}}_{G'} \cdot \underbrace{\frac{\partial(x-ct)}{\partial x}}_1 \\ &= F'(x+ct) + G'(x-ct)\end{aligned}\quad (2.4a)$$

$$\frac{\partial^2 u}{\partial x^2} = F''(x+ct) + G''(x-ct) \quad (2.4b)$$

and substitute into the scalar wave equation (2.1)

$$c^2 F''(x+ct) + c^2 G''(x-ct) = c^2 [F''(x+ct) + G''(x-ct)] \quad (2.5)$$

This is an identity regardless of the choice of F and G .

F and G are known as propagating-wave solutions. For example, consider a “snapshot” of $F(x+ct)$ taken at time t_0 . After Δt seconds have passed, the wave solution F has moved to the left (in the $-x$ direction) by $c\Delta t$ meters. Why? Well, the time part of the argument of F has *increased* by $c\Delta t$. Therefore the space part of the argument has to *decrease* by $c\Delta t$ to obtain the previous wave function value. The converse is true for $G(x-ct)$, a rightward traveling wave. We see that c represents the speed of wave propagation in the $+x$ or $-x$ direction.

2.3 DISPERSION RELATION

In this section we introduce the concept of dispersion. Dispersion is defined as the variation of a propagating wave’s wavelength λ with frequency f . For convenience, dispersion is also frequently represented as the variation of the propagating wave’s wavenumber $k = 2\pi/\lambda$ with angular frequency $\omega = 2\pi f$.

The simplest example of a dispersion relation results from the one-dimensional scalar wave equation itself. Consider a continuous sinusoidal-traveling-wave solution of (2.1), as written in phasor form:

$$u(x, t) = e^{j(\omega t - kx)} \quad (2.6)$$

where ω and k are defined above and $j = \sqrt{-1}$. Substituting into (2.1), we obtain

$$(j\omega)^2 e^{j(\omega t - kx)} = c^2 (-jk)^2 e^{j(\omega t - kx)} \quad (2.7a)$$

Now factoring out the complex exponential term common to both sides yields

$$\begin{aligned} -\omega^2 &= c^2 \cdot (-k^2) \\ \omega^2 &= c^2 k^2 \quad \rightarrow \quad k = \pm \omega / c \end{aligned} \quad (2.7b)$$

In fact, (2.7b) is the dispersion relation for the one-dimensional scalar wave equation (2.1). This relation is very simple, stating that the wavenumber is linearly proportional to the sinusoidal frequency. The plus sign designates $+x$ -directed wave propagation, while the minus sign designates $-x$ -directed propagation.

From (2.7b) we can obtain an expression for the wave *phase velocity*, classically defined as $v_p = \omega/k$:

$$v_p = \frac{\omega}{k} = \pm c \quad (2.8)$$

The phase velocity is seen to be $\pm c$, a constant regardless of frequency. Propagating waves having a dispersion relation of the form of (2.7b), resulting in the constant phase velocity of (2.8), are said to be *dispersionless*. In effect, their waveshape remains unchanged after arbitrarily large propagation distances for arbitrary modulation envelopes or pulse shapes.

Further, by considering the angular frequency to be a function of the wavenumber (i.e., $\omega = \omega(k)$), we can differentiate (2.7b) with respect to k to obtain the wave *group velocity*, classically defined as $v_g = d\omega/dk$. This yields

$$\begin{aligned} 2\omega \frac{d\omega}{dk} &= c^2 \cdot (2k) \\ v_g &= \frac{d\omega}{dk} = \frac{2c^2 k}{2\omega} = \frac{c^2}{\omega} \cdot \left(\pm \frac{\omega}{c} \right) = \pm c \end{aligned} \quad (2.9)$$

The group velocity for this case is also seen to be $\pm c$, independent of frequency.

2.4 FINITE DIFFERENCES

Consider a Taylor's series expansion of $u(x, t_n)$ about the space point x_i to the space point $x_i + \Delta x$, keeping time fixed at t_n :

$$\begin{aligned} u(x_i + \Delta x) \Big|_{t_n} &= u|_{x_i, t_n} + \Delta x \cdot \frac{\partial u}{\partial x} \Big|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \cdot \frac{\partial^2 u}{\partial x^2} \Big|_{x_i, t_n} + \frac{(\Delta x)^3}{6} \cdot \frac{\partial^3 u}{\partial x^3} \Big|_{x_i, t_n} \\ &\quad + \frac{(\Delta x)^4}{24} \cdot \frac{\partial^4 u}{\partial x^4} \Big|_{\xi_1, t_n} \end{aligned} \quad (2.10a)$$

The last term is the remainder, or error, term. Here, ξ_1 is a space point located somewhere in the interval $(x_i, x_i + \Delta x)$. Similarly, consider the Taylor's series expansion to the space point $x_i - \Delta x$, again keeping time fixed at t_n :

$$\begin{aligned} u(x_i - \Delta x) \Big|_{t_n} &= u|_{x_i, t_n} - \Delta x \cdot \frac{\partial u}{\partial x} \Big|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \cdot \frac{\partial^2 u}{\partial x^2} \Big|_{x_i, t_n} - \frac{(\Delta x)^3}{6} \cdot \frac{\partial^3 u}{\partial x^3} \Big|_{x_i, t_n} \\ &\quad + \frac{(\Delta x)^4}{24} \cdot \frac{\partial^4 u}{\partial x^4} \Big|_{\xi_2, t_n} \end{aligned} \quad (2.10b)$$

In the remainder term, ξ_2 is a space point located somewhere in the interval $(x_i, x_i - \Delta x)$. Now, adding (2.10a) and (2.10b), we obtain

$$u(x_i + \Delta x) \Big|_{t_n} + u(x_i - \Delta x) \Big|_{t_n} = 2u|_{x_i, t_n} + (\Delta x)^2 \cdot \frac{\partial^2 u}{\partial x^2} \Big|_{x_i, t_n} + \frac{(\Delta x)^4}{12} \cdot \frac{\partial^4 u}{\partial x^4} \Big|_{\xi_3, t_n} \quad (2.11)$$

Here, by the mean value theorem, ξ_3 is a space point located somewhere in the interval $(x_i - \Delta x, x_i + \Delta x)$. Rearranging terms, we obtain

$$\frac{\partial^2 u}{\partial x^2} \Big|_{x_i, t_n} = \left[\frac{u(x_i + \Delta x) - 2u(x_i) + u(x_i - \Delta x)}{(\Delta x)^2} \right]_{t_n} + O[(\Delta x)^2] \quad (2.12)$$

where $O[(\Delta x)^2]$ is a shorthand notation for the remainder term, which approaches zero as the square of the space increment. Equation (2.12) is commonly referred to as a second-order accurate, central-difference approximation to the second partial space derivative of u . For convenience, we will adopt a subscript i for the space position and a superscript n for the time observation point. Equation (2.12) is rewritten as:

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{x_i, t_n} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} + O[(\Delta x)^2] \quad (2.13)$$

In (2.13) and all subsequent finite-difference expressions, it is understood that u_i^n denotes a field quantity calculated at the space point $x_i = i\Delta x$ and time point $t_n = n\Delta t$.

For the second partial time derivative, we keep x_i fixed and expand u in forward and backward Taylor's series in time. By analogy with (2.13), we obtain a second-order accurate, central-difference approximation to the second partial time derivative of u :

$$\left. \frac{\partial^2 u}{\partial t^2} \right|_{x_i, t_n} = \frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{(\Delta t)^2} + O[(\Delta t)^2] \quad (2.14)$$

2.5 FINITE-DIFFERENCE APPROXIMATION OF THE SCALAR WAVE EQUATION

Substituting the two central-difference expressions of (2.13) and (2.14) into the one-dimensional scalar wave equation (2.1), we have

$$\frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{(\Delta t)^2} + O[(\Delta t)^2] = c^2 \left\{ \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} + O[(\Delta x)^2] \right\} \quad (2.15)$$

This approximation to the scalar wave equation has a second-order accuracy in both space and time, $O[(\Delta t)^2 + (\Delta x)^2]$, which is denoted succinctly as “a (2, 2) accurate solution.” (See this chapter’s appendix for additional discussion of the order of accuracy.) Neglecting the Taylor series remainder terms and solving for the latest value of u at grid point i , we obtain

$$u_i^{n+1} \cong (c\Delta t)^2 \left[\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} \right] + 2u_i^n - u_i^{n-1} \quad (2.16)$$

This is a *fully explicit* second-order accurate expression for u_i^{n+1} in that all wave quantities on the right-hand side are known; that is, they were obtained during the previous time steps, n and $n-1$, and then stored in the computer memory. No simultaneous equation solution is needed. Upon performing (2.16) for all space points of interest, yielding the complete set of u_i^{n+1} , the process can begin again to obtain u_i^{n+2} . Repetition of (2.16) over the problem space constitutes the numerical FDTD solution of the scalar wave equation.

Of particular interest will be the case $c\Delta t = \Delta x$. The properties of the resulting finite-difference expression are so remarkable that henceforth we will refer to this situation as the *magic time-step*. For this case, (2.16) reduces to simply

$$\begin{aligned} u_i^{n+1} &= (u_{i+1}^n - 2u_i^n + u_{i-1}^n) + 2u_i^n - u_i^{n-1} \\ &= u_{i+1}^n + u_{i-1}^n - u_i^{n-1} \end{aligned} \quad (2.17)$$

Note that (2.17) denotes an exact equality of the left- and right-hand sides, not an approximation as shown in (2.16). This is *not* an error. In fact, for $c\Delta t = \Delta x$, we can show that the solution to the numerical finite-difference equation is an *exact* solution to the original differential wave equation (2.1) despite the Taylor's series approximations. This proof is carried out next.

Consider again $u(x, t) = F(x + ct) + G(x - ct)$, the exact propagating-wave solutions to the one-dimensional scalar wave equation (2.1). In our notation, u evaluated at space point x_i and time point t_n is given by $u_i^n = F(x_i + ct_n) + G(x_i - ct_n)$, where $x_i = i\Delta x$ and $t_n = n\Delta t$. Now assume that these exact propagating-wave data are somehow available and stored in computer memory. Further, assume that we perform the simple algorithm of (2.17) upon these exact data, representing a time advance of Δt :

$$\begin{aligned} u_i^{n+1} &= u_{i+1}^n + u_{i-1}^n - u_i^{n-1} \\ \downarrow &\quad \downarrow &\quad \downarrow &\quad \downarrow \\ \left[\begin{array}{l} F(x_i + ct_{n+1}) \\ + G(x_i - ct_{n+1}) \end{array} \right] &= \left[\begin{array}{l} F(x_{i+1} + ct_n) \\ + G(x_{i+1} - ct_n) \end{array} \right] + \left[\begin{array}{l} F(x_{i-1} + ct_n) \\ + G(x_{i-1} - ct_n) \end{array} \right] - \left[\begin{array}{l} F(x_i + ct_{n-1}) \\ + G(x_i - ct_{n-1}) \end{array} \right] \end{aligned} \quad (2.18a)$$

Initially, we see that the algorithm yields six terms, three rightward and three leftward propagating waves. However, some cancellation of these terms will soon be apparent. Expanding the right-hand side (RHS) of (2.18a), we obtain

$$\begin{aligned} \text{RHS} &= \left\{ \begin{array}{l} F[(i+1)\Delta x + cn\Delta t] \\ + G[(i+1)\Delta x - cn\Delta t] \end{array} \right\} + \left\{ \begin{array}{l} F[(i-1)\Delta x + cn\Delta t] \\ + G[(i-1)\Delta x - cn\Delta t] \end{array} \right\} \\ &- \left\{ \begin{array}{l} F[i\Delta x + c(n-1)\Delta t] \\ + G[i\Delta x - c(n-1)\Delta t] \end{array} \right\} \end{aligned} \quad (2.18b)$$

From the magic time step, we let $c\Delta t = \Delta x$ in each argument of F and G in (2.18b):

$$\begin{aligned}
\text{RHS} &= \left\{ \begin{array}{l} F[(i+1)\Delta x + n\Delta x] \\ + G[(i+1)\Delta x - n\Delta x] \end{array} \right\} + \left\{ \begin{array}{l} F[(i-1)\Delta x + n\Delta x] \\ + G[(i-1)\Delta x - n\Delta x] \end{array} \right\} - \left\{ \begin{array}{l} F[i\Delta x + (n-1)\Delta x] \\ + G[i\Delta x - (n-1)\Delta x] \end{array} \right\} \\
&= \left\{ \begin{array}{l} F[(i+1+n)\Delta x] \\ + G[(i+1-n)\Delta x] \end{array} \right\} + \left\{ \begin{array}{l} F[(i-1+n)\Delta x] \\ + G[(i-1-n)\Delta x] \end{array} \right\} - \left\{ \begin{array}{l} F[(i+n-1)\Delta x] \\ + G[(i-n+1)\Delta x] \end{array} \right\}
\end{aligned} \tag{2.18c}$$

It is clear that the G term in the first bracket cancels the G term in the third bracket, and the F term in the second bracket cancels the F term in the third bracket. This leaves only two terms for the right-hand side:

$$\text{RHS} = F[(i+1+n)\Delta x] + G[(i-1-n)\Delta x] \tag{2.18d}$$

However, these two terms represent the *same* two propagating waves that we started out with, only shifted in space to the left and to the right by the distance $c\Delta t = \Delta x$, the distance moved by each wave traveling at the speed c over one time step Δt , as prescribed by the original wave equation (2.1). We can see this simply by evaluating the initially assumed exact wave solution u at the original space point x_i but with the time advanced to $t_n + \Delta t$:

$$\begin{aligned}
u_i^{n+1} &= F(x_i + ct_{n+1}) + G(x_i - ct_{n+1}) \\
&= F[i\Delta x + c(n+1)\Delta t] + G[i\Delta x - c(n+1)\Delta t]
\end{aligned} \tag{2.19a}$$

Again replacing $c\Delta t$ with Δx in the arguments of F and G , we obtain

$$u_i^{n+1} = F[(i+n+1)\Delta x] + G[(i-n-1)\Delta x] \tag{2.19b}$$

We note that this is exactly the right-hand side of (2.18d). Therefore, the magic time-step algorithm of (2.17) results in the assumed exact initial data for the propagating waves being transformed to new exact data for the waves, with the proper physics of propagation incorporated. If desired, we can then apply (2.17) again, but now to the new wave data obtained at time step $n+1$. In this manner, we can proceed to time-step $n+2$. By induction, it is clear that we can proceed any number of time-steps, or iterations of the algorithm, *always calculating exact propagating-wave data*.

This is a startling result because, apparently, the approximations inherent in the Taylor's series derivation of the space and time derivatives cancel out for the magic time-step, yielding a numerical finite-difference solution for the continuous one-dimensional wave equation that is exact. In the next section we shall explore the theoretical background behind this in more depth.

2.6 NUMERICAL DISPERSION RELATION

The procedure used in Section 2.3 can be applied to obtain the *numerical dispersion* relation of the finite-difference approximation of the one-dimensional scalar wave equation given by (2.16). Consider the sinusoidal wave of (2.6) of angular frequency ω discretely sampled in space and time at (x_i, t_n) in a finite-difference grid. Let $\tilde{k} = \tilde{k}_{real} + j\tilde{k}_{imag}$ be the possibly complex-valued wavenumber of this “numerical” wave. Then, (2.6) becomes

$$u_i^n = e^{j(\omega n\Delta t - \tilde{k} i\Delta x)} = e^{j[\omega n\Delta t - (\tilde{k}_{real} + j\tilde{k}_{imag}) i\Delta x]} = e^{\tilde{k}_{imag} i\Delta x} e^{j(\omega n\Delta t - \tilde{k}_{real} i\Delta x)} \quad (2.20)$$

In general, \tilde{k} differs from k , the corresponding physical wavenumber. This difference, called the numerical-dispersion artifact, may give rise to numerical wave amplitudes and velocities that depart from the exact values. For later reference, we note that (2.20) permits either a constant wave amplitude with spatial position ($\tilde{k}_{imag} = 0$); an exponentially decreasing amplitude with spatial position ($\tilde{k}_{imag} < 0$); or an exponentially increasing amplitude with spatial position ($\tilde{k}_{imag} > 0$).

We now substitute into (2.16) the numerical traveling wave of (2.20). This yields

$$\begin{aligned} e^{j[\omega(n+1)\Delta t - \tilde{k} i\Delta x]} &= \left(\frac{c\Delta t}{\Delta x}\right)^2 \left\{ e^{j[\omega n\Delta t - \tilde{k}(i+1)\Delta x]} - 2e^{j[\omega n\Delta t - \tilde{k} i\Delta x]} + e^{j[\omega n\Delta t - \tilde{k}(i-1)\Delta x]} \right\} \\ &\quad + 2e^{j[\omega n\Delta t - \tilde{k} i\Delta x]} - e^{j[\omega(n-1)\Delta t - \tilde{k} i\Delta x]} \end{aligned} \quad (2.21a)$$

After factoring out $\exp j(\omega n\Delta t - \tilde{k} i\Delta x)$ on both sides, we obtain

$$e^{j\omega\Delta t} = \left(\frac{c\Delta t}{\Delta x}\right)^2 \cdot (e^{-j\tilde{k}\Delta x} - 2 + e^{j\tilde{k}\Delta x}) + 2 - e^{-j\omega\Delta t} \quad (2.21b)$$

Grouping the time and space exponential terms and dividing both sides by 2 yields

$$\frac{e^{j\omega\Delta t} + e^{-j\omega\Delta t}}{2} = \left(\frac{c\Delta t}{\Delta x}\right)^2 \left(\frac{e^{j\tilde{k}\Delta x} + e^{-j\tilde{k}\Delta x}}{2} - 1 \right) + 1 \quad (2.21c)$$

Upon applying Euler’s identity to the complex exponentials, we obtain the numerical dispersion relation corresponding to (2.16) expressed in two equivalent, useful forms:

$$\cos(\omega \Delta t) = \left(\frac{c \Delta t}{\Delta x} \right)^2 [\cos(\tilde{k} \Delta x) - 1] + 1 \quad (2.22)$$

$$\tilde{k} = \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + \left(\frac{\Delta x}{c \Delta t} \right)^2 [\cos(\omega \Delta t) - 1] \right\} \quad (2.23)$$

We see that (2.22) and (2.23) are much more complicated relations between ω and \tilde{k} than the straightforward algebraic expression of (2.7b), the dispersion relation for the continuous one-dimensional wave equation. Nevertheless, we can use (2.22) and (2.23) to obtain information about the nature of the grid's numerical traveling waves represented by (2.20). We shall consider three cases as follows.

2.6.1 Case 1: Very Fine Sampling in Time and Space ($\Delta t \rightarrow 0, \Delta x \rightarrow 0$)

Here, we start with (2.23) and assume a fine-time-sampling condition such that the $\omega \Delta t$ argument of the cosine approaches zero. This permits applying a two-term Taylor's series expansion to this cosine:

$$\begin{aligned} \tilde{k} &= \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + \left(\frac{\Delta x}{c \Delta t} \right)^2 \left[1 - \frac{(\omega \Delta t)^2}{2} - 1 \right] \right\} \\ &= \frac{1}{\Delta x} \cos^{-1} \left[1 - \frac{1}{2} \left(\frac{\omega}{c} \right)^2 (\Delta x)^2 \right] = \frac{1}{\Delta x} \cos^{-1} \left[1 - \frac{1}{2} (k \Delta x)^2 \right] \end{aligned} \quad (2.24a)$$

where the free-space wavenumber is $k = \omega/c$. Next, we note that $k \Delta x$ approaches zero under the assumed fine-space-sampling condition. This permits applying a two-term Taylor's series expansion to the \cos^{-1} function, which yields our final result:

$$\tilde{k} = \frac{1}{\Delta x} \cdot (k \Delta x) = k \quad (2.24b)$$

We see that the numerical wavenumber \tilde{k} in (2.24b) is exactly the free-space wavenumber. Since the latter provides for phase velocity v_p and group velocity v_g equal to c regardless of frequency (i.e., dispersionless wave propagation), we infer that the numerical solution for this case is dispersionless as well. Thus, we have a satisfying intuitive result: in the limit as the space and time increments of our finite-difference approximations go to zero, the numerical solution becomes exact.

2.6.2 Case 2: Magic Time-Step ($c \Delta t = \Delta x$)

In this case, we substitute the magic-time-step relation into (2.23). This yields

$$\tilde{k} = \frac{1}{c\Delta t} \cos^{-1}[1 + \cos(\omega\Delta t) - 1] \quad (2.25a)$$

Performing the indicated cancellation, we obtain

$$\tilde{k} = \frac{1}{c\Delta t} \cos^{-1}[\cos(\omega\Delta t)] = \frac{\omega\Delta t}{c\Delta t} = k \quad (2.25b)$$

Just as in Case 1, the numerical wavenumber \tilde{k} is exactly the free-space wavenumber, and the numerical solution is exact. However, unlike Case 1, this is nonintuitive because it implies that the numerical solution is exact regardless of the choice of the space and time increments (fine *or* coarse). The space and time increments need only be scaled by the free-space wave-propagation velocity c .

2.6.3 Case 3: Dispersive Wave Propagation

This is the general solution of (2.23) where finite numerical dispersion errors can exist. Before wrestling with the mathematics of the general solution, it is helpful to obtain sample values of \tilde{k} and numerical phase velocity. These results show that the phase velocity of a sinusoidal numerical wave within the grid is determined by the grid's sampling resolution relative to the free-space wavelength.

Sample Values of the Numerical Wavenumber and Phase Velocity

We consider first a choice of space and time increments such that $c\Delta t = \Delta x/2$ and $\Delta x = \lambda_0/10$, where λ_0 is the free-space wavelength of a continuous propagating sinusoidal wave of angular frequency ω . From (2.23), we obtain

$$\begin{aligned} \tilde{k} &= \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + \left(\frac{\Delta x}{\Delta x/2} \right)^2 \left[\cos \left(\frac{\omega \Delta x}{2c} \right) - 1 \right] \right\} \\ &= \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + 4 \left[\cos \left(\frac{k \Delta x}{2} \right) - 1 \right] \right\} \end{aligned} \quad (2.26a)$$

where again $k = \omega/c$. Now further noting that $k = 2\pi/\lambda_0$ and that we have assumed that $\Delta x/\lambda_0 = 0.1$, we proceed:

$$\begin{aligned} \tilde{k} &= \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + 4 \left[\cos \left(\frac{2\pi}{\lambda_0} \cdot \frac{\Delta x}{2} \right) - 1 \right] \right\} \\ &= \frac{1}{\Delta x} \cos^{-1}(0.8042) = \frac{0.63642}{\Delta x} \end{aligned} \quad (2.26b)$$

Defining a numerical phase velocity \tilde{v}_p by analogy to the phase velocity for the continuous-wave-equation case (see (2.8)):

$$\tilde{v}_p = \frac{\omega}{\tilde{k}} \quad (2.27a)$$

we have

$$\tilde{v}_p = \frac{2\pi f}{(0.63642 / \Delta x)} = \frac{2\pi(c/\lambda_0)\Delta x}{0.63642} = \frac{2\pi \cdot (0.1)}{0.63642} c = 0.9873c \quad (2.27b)$$

The numerical phase velocity for this example is seen to be 1.27% less than the free-space phase velocity c . That is, for a physical wave propagating over a distance equivalent to $10\lambda_0$ (100 space cells), the numerical analog would propagate only 98.73 cells. At the leading edge of the propagating wave, this represents a numerical phase error of $[(100 - 98.73)/10] \cdot 360^\circ$, or 45.72° .

We next repeat the above steps for Δx halved, i.e., a grid space resolution of $\Delta x = \lambda_0/20$. This yields $\tilde{k} = 0.31514 / \Delta x$ and $\tilde{v}_p = 0.99689c$, implying a numerical phase-velocity error of -0.31%. We see that reducing Δx by 2:1 cuts the numerical phase-velocity error by about 4:1, indicative of the second-order accuracy of the finite-difference algorithm that we are using. Now for a physical wave propagating over the same $10\lambda_0$ distance considered above (here equivalent to 200 space cells), the numerical analog would propagate 199.378 cells. At the leading edge of the propagating wave, this represents a numerical phase error of $[(200 - 199.378)/20] \cdot 360^\circ$, or 11.196° , again reduced by a factor of about 4:1.

General Solution

Having established that the grid space resolution directly impacts the numerical phase velocity, it is very useful to solve (2.23) for the general case of Δt and Δx . Following the notation of [1], we let

$$S = \frac{c \Delta t}{\Delta x} \quad (2.28a)$$

be the numerical stability factor (or Courant number), to be discussed later in this chapter. Further, we define

$$N_\lambda = \frac{\lambda_0}{\Delta x} \quad (2.28b)$$

to be the grid sampling resolution in space cells per free-space wavelength. Note that N_λ is defined in terms of resolving the continuous-world wavelength λ_0 and *not* the numerical wavelength $\tilde{\lambda}$ in the FDTD grid. Then, (2.23) can be written as

$$\begin{aligned}\tilde{k} &= \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + \left(\frac{1}{S} \right)^2 \left[\cos \left(\frac{2\pi S}{N_\lambda} \right) - 1 \right] \right\} \\ &= \frac{1}{\Delta x} \cos^{-1}(\zeta) = \frac{1}{\Delta x} \left[\frac{\pi}{2} - \sin^{-1}(\zeta) \right]\end{aligned}\quad (2.29a)$$

where

$$\zeta = 1 + \left(\frac{1}{S} \right)^2 \left[\cos \left(\frac{2\pi S}{N_\lambda} \right) - 1 \right] \quad (2.29b)$$

As discussed in [1], caution must be exercised in evaluating numerical dispersion relations such as (2.29) since it is possible to choose S and N_λ such that \tilde{k} is complex. In the specific case of (2.29), it can be shown that the transition between real and complex values of \tilde{k} occurs when $\zeta = -1$. Solving for N_λ at this transition results in

$$N_\lambda|_{\text{transition}} = \frac{2\pi S}{\cos^{-1}(1 - 2S^2)} \quad (2.30)$$

For grid sampling finer than this value, i.e. $N_\lambda > N_\lambda|_{\text{transition}}$, \tilde{k} is a real number and the numerical wave undergoes no attenuation while propagating in the grid. Here, the numerical phase velocity \tilde{v}_p is less than the free-space phase velocity c , as shown in the example resulting in (2.27b). For coarser grid sampling $N_\lambda < N_\lambda|_{\text{transition}}$, \tilde{k} is a complex number and the numerical wave undergoes a nonphysical exponential decay while propagating [1]. Further, in this coarse-resolution regime, \tilde{v}_p can exceed c [1].

We now illustrate how \tilde{k} and \tilde{v}_p varies with grid sampling N_λ both above and below the transition between real and complex numerical wavenumbers. For clarity, we choose a specific example wherein $S = 0.5$. This yields $N_\lambda|_{\text{transition}} = 3$ from (2.30).

Real-Numerical-Wavenumber Regime

For $N_\lambda \geq 3$ we have from (2.29)

$$\tilde{k}_{\text{real}} = \frac{1}{\Delta x} \cos^{-1} \left\{ 1 + 4 \left[\cos \left(\frac{\pi}{N_\lambda} \right) - 1 \right] \right\}; \quad \tilde{k}_{\text{imag}} = 0 \quad (2.31)$$

The numerical phase velocity is given by

$$\tilde{v}_p = \frac{\omega}{\tilde{k}_{\text{real}}} = \frac{2\pi f \Delta x}{\cos^{-1} \left\{ 1 + 4 \left[\cos \left(\frac{\pi}{N_\lambda} \right) - 1 \right] \right\}} \quad (2.32a)$$

Since

$$2\pi f \Delta x = 2\pi f \left(\frac{\lambda_0}{N_\lambda} \right) = \left(\frac{2\pi}{N_\lambda} \right) c \quad (2.32b)$$

we can rewrite (2.32a) as

$$\tilde{v}_p = \frac{2\pi}{N_\lambda \cos^{-1} \left\{ 1 + 4 \left[\cos \left(\frac{\pi}{N_\lambda} \right) - 1 \right] \right\}} c \quad (2.32c)$$

From (2.20), we note that the wave-amplitude multiplier per grid cell of propagation is:

$$e^{\tilde{k}_{\text{imag}} \Delta x} \equiv e^{-\alpha \Delta x} = e^0 = 1 \quad (2.33)$$

Thus, there is a constant wave amplitude with spatial position for this range of N_λ .

Complex-Numerical-Wavenumber Regime

For $N_\lambda < 3$, we observe that $\zeta < -1$ in (2.29). The following relation for the complex-valued arc-sine function [1, 2] becomes useful:

$$\sin^{-1}(\zeta) = -j \ln \left(j\zeta + \sqrt{1 - \zeta^2} \right) \quad (2.34)$$

Substituting (2.34) into (2.29) yields

$$\begin{aligned} \tilde{k} &= \frac{1}{\Delta x} \left[\frac{\pi}{2} + j \ln \left(j\zeta + \sqrt{1 - \zeta^2} \right) \right] \\ &= \frac{1}{\Delta x} \left[\frac{\pi}{2} + j \ln \left(j\zeta + j\sqrt{\zeta^2 - 1} \right) \right] \end{aligned} \quad (2.35a)$$

Factoring out $j = e^{+j\pi/2}$ in the argument of the natural logarithm, we obtain

$$\begin{aligned} \tilde{k} &= \frac{1}{\Delta x} \left\{ \frac{\pi}{2} + j \ln \left[\left(\zeta + \sqrt{\zeta^2 - 1} \right) e^{+j\pi/2} \right] \right\} \\ &= \frac{1}{\Delta x} \left\{ \frac{\pi}{2} + j \ln \left[\left(-\zeta - \sqrt{\zeta^2 - 1} \right) e^{-j\pi/2} \right] \right\} \end{aligned} \quad (2.35b)$$

Upon taking the natural logarithm, we further obtain

$$\begin{aligned}\tilde{k} &= \frac{1}{\Delta x} \left\{ \frac{\pi}{2} + j \left[\ln \left(-\zeta - \sqrt{\zeta^2 - 1} \right) - j \frac{\pi}{2} \right] \right\} \\ &= \frac{1}{\Delta x} \left\{ \pi + j \ln \left(-\zeta - \sqrt{\zeta^2 - 1} \right) \right\}\end{aligned}\quad (2.35c)$$

This yields

$$\tilde{k}_{real} = \frac{\pi}{\Delta x}; \quad \tilde{k}_{imag} = \frac{1}{\Delta x} \ln \left(-\zeta - \sqrt{\zeta^2 - 1} \right) \quad (2.36)$$

Following [1], the numerical phase velocity for this case is

$$\begin{aligned}\tilde{v}_p &= \frac{\omega}{\tilde{k}_{real}} = \frac{\omega}{(\pi / \Delta x)} = \frac{2\pi f \Delta x}{\pi} \\ &= \frac{2f\lambda_0}{N_\lambda} = \frac{2}{N_\lambda} c\end{aligned}\quad (2.37a)$$

and from (2.20) the wave-amplitude multiplier per grid cell of propagation is

$$e^{\tilde{k}_{imag} \Delta x} \equiv e^{-\alpha \Delta x} = e^{\ln(-\zeta - \sqrt{\zeta^2 - 1})} = -\zeta - \sqrt{\zeta^2 - 1} \quad (2.37b)$$

Since $\zeta < -1$, the numerical wave amplitude decays exponentially with spatial position.

Let us now consider the possibility of \tilde{v}_p exceeding c in this situation [1]. Nyquist theory states that any physical or numerical process that obtains samples of a time waveform every Δt seconds can reproduce the original waveform (without aliasing) for spectral content up to $f_{max} = 1/(2\Delta t)$. In the present case, the corresponding minimum free-space wavelength that can be sampled without aliasing is therefore

$$\lambda_{0,min} = c/f_{max} = 2c\Delta t \quad (2.38a)$$

For a particular choice of grid Δx , the corresponding minimum spatial sampling is

$$N_{\lambda,min} = \lambda_{0,min}/\Delta x = 2c\Delta t/\Delta x = 2S \quad (2.38b)$$

Since (as will be shown later in this chapter) $S \leq 1$ to insure numerical stability of the algorithm, we have

$$N_{\lambda,min} = 2S \leq 2 \quad (2.38c)$$

Then from (2.37a) the maximum numerical phase velocity is given by

$$\tilde{v}_{p,\max} = \frac{2}{N_{\lambda,\min}} c = \frac{2}{2S} c = \frac{1}{S} c \geq c \quad (2.39a)$$

We see that \tilde{v}_p can exceed c for this case. In fact, substituting S from (2.28a) into (2.39a), we obtain the maximum possible numerical phase velocity in the grid [1]:

$$\tilde{v}_{p,\max} = \frac{1}{S} c = \frac{\Delta x}{c\Delta t} c = \frac{\Delta x}{\Delta t} \quad (2.39b)$$

This relation tells us that in one Δt , a numerical value can propagate at most one Δx . This is intuitively correct given the local nature of the spatial difference used in the field-updating algorithm. That is, a field point more than one Δx away from a source point that undergoes a sudden change cannot possibly “feel” the effect of that change during the next Δt . Note that $\tilde{v}_{p,\max}$ is independent of material parameters and is an inherent property of the grid and its method of obtaining space derivatives.

2.6.4 Example of Calculation of Numerical Phase Velocity and Attenuation

This section provides an example of a detailed calculation of numerical phase velocity and attenuation based upon the dispersion analysis of Section 2.6.3. Graphs are provided that depict the variation of the numerical phase velocity and attenuation as a function of the grid sampling density.

Fig. 2.1 graphs \tilde{v}_p/c , the numerical phase velocity normalized to the free-space speed of light, and $\alpha\Delta x$, the exponential attenuation constant per grid cell, versus the grid sampling density N_λ (given in points per free-space wavelength λ_0). A stability factor (Courant number) $S = 0.5$ is assumed. The data shown in this figure are obtained from (2.32c) and (2.33) for the regime $N_\lambda \geq 3$, and from (2.37a) and (2.37b) for the regime $N_\lambda < 3$. We see from Fig. 2.1 that, at $N_\lambda = 3$ (the transition point between real and complex values of \tilde{k}), a minimum value of $\tilde{v}_p = (2/3)c$ is reached. This point is also the onset of exponential attenuation of numerical waves in the grid. As N_λ is reduced below 3, \tilde{v}_p increases inversely with N_λ . Eventually, \tilde{v}_p exceeds c for $N_\lambda < 2$, and reaches a limiting velocity of $2c$ as $N_\lambda \rightarrow 1$. In this limit, as well, the attenuation constant approaches a value of 2.639 nepers/cell. It appears that very coarsely-resolved wave modes in the grid can propagate at superluminal speeds [1]. However, these modes are rapidly attenuated [1].

Fig. 2.2 graphs the percent error in the numerical phase velocity relative to the free-space speed of light in the grid-sampling density regime $N_\lambda \geq 3$. Similar to Fig. 2.1, a stability factor (Courant number) $S = 0.5$ is assumed. As $N_\lambda \gg 10$, we see from Fig. 2.2 that the numerical phase-velocity error diminishes as the inverse square of N_λ . This is indicative of the second-order-accurate nature of the algorithm.

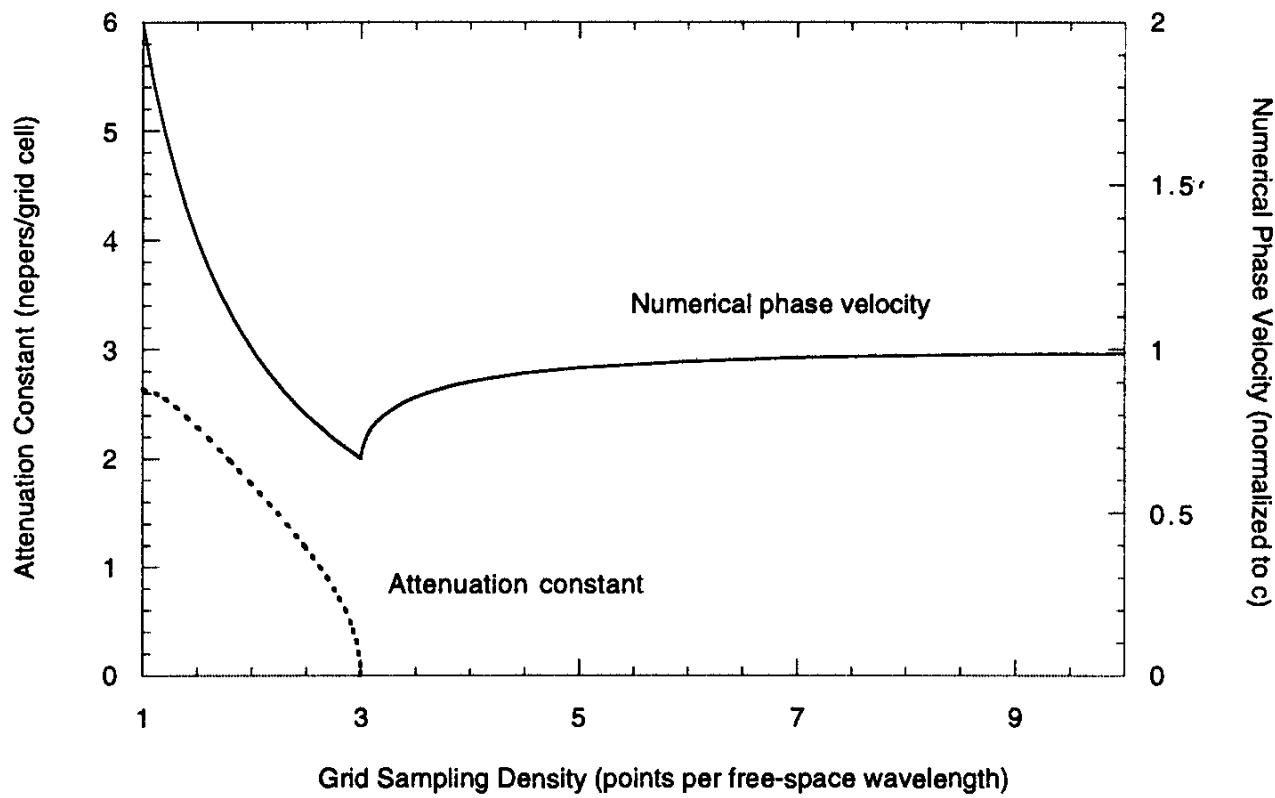


Fig. 2.1 Variation of the normalized numerical phase velocity \tilde{v}_p/c and attenuation per grid cell $\alpha\Delta x$ as a function of the grid sampling density ($1 \leq N_\lambda \leq 10$) for a Courant stability factor $S = 0.5$.

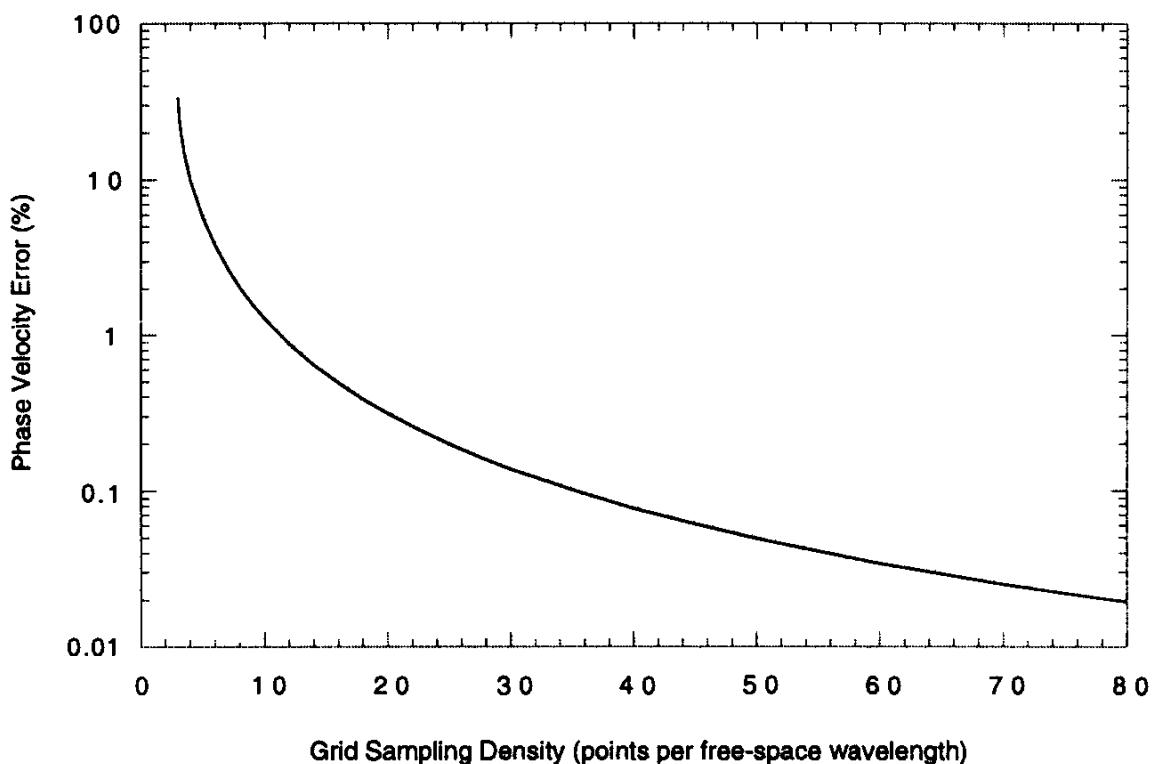


Fig. 2.2 Percent numerical phase-velocity error relative to the free-space speed of light as a function of the grid sampling density ($3 \leq N_\lambda \leq 80$) for a Courant stability factor $S = 0.5$.

2.6.5 Examples of Calculations of Pulse Propagation

Fig. 2.3(a) graphs examples of the calculated propagation of a 40-cell-wide rectangular pulse in free space for two cases of the Courant stability factor: $S = 1$ (i.e., Δt is equal to the magic time step); and $S = 0.99$ (i.e., Δt is just 1% below the magic time-step). To permit a direct comparison of these results, both “snapshots” are taken at the same absolute time after the onset of time-stepping. There are three key observations:

1. When $S = 1$, the rectangular shape and spatial width of the pulse are completely preserved. For this case, the abrupt step discontinuities of the propagating pulse are modeled perfectly. In fact, this is expected since $\tilde{v}_p \equiv c$ for all numerical modes in the grid under magic-time-step conditions.
2. When $S = 0.99$, the step discontinuities at the leading and trailing edges of the pulse generate appreciable “ringing” which lags behind the corresponding discontinuity. In fact, this ringing represents time-retarded propagation (i.e., $\tilde{v}_p < c$) of the sparsely sampled, high-frequency spectral content generated by the step discontinuities.
3. When $S = 0.99$, a weak superluminal response (i.e., $\tilde{v}_p > c$) propagates just ahead of the leading edge of the pulse. This is in accordance with the dispersive propagation theory developed earlier.

Fig. 2.3(b) repeats the examples of Fig. 2.3(a), but for the Courant stability factors $S = 1$ and $S = 0.5$ (i.e., Δt is set at 50% below the magic time-step). Again, to permit a direct comparison of the results, both “snapshots” are taken at the same absolute time after the onset of time-stepping. From this figure, we see that the duration and periodicity of the ringing is significantly greater than that for the $S = 0.99$ case of Fig. 2.3(a). Further, the superluminal response located ahead of the pulse leading edge is more pronounced and less damped.

Figs. 2.4(a) and 2.4(b) repeat the examples of Figs. 2.3(a) and 2.3(b), but for a smooth Gaussian pulse having a 40-grid-cell spatial width between $1/e$ points. Again, to permit a direct comparison of the results, all “snapshots” are taken at the same absolute time after the onset of time-stepping. We see that the propagating Gaussian pulse undergoes much less distortion than the rectangular pulse. The calculated pulse propagation for $S = 0.99$ shows no observable difference (at the scale of Fig. 2.4(a)) relative to the perfect propagation case of $S = 1$. Even for $S = 0.5$, the calculated pulse propagation shows only a slight retardation relative to the exact solution, as expected because $\tilde{v}_p < c$ for virtually all modes in the grid. Further, there is no observable superluminal precursor. All of these phenomena are due to the fact that, for this case, virtually the entire spatial spectrum of propagating wavelengths within the grid is well resolved by the grid’s sampling process. As a result, almost all numerical phase-velocity errors relative to the free-space speed of light are well below 1%, as seen from Fig. 2.2. This allows the Gaussian pulse to “hold together” while propagating over significant distances within the grid.

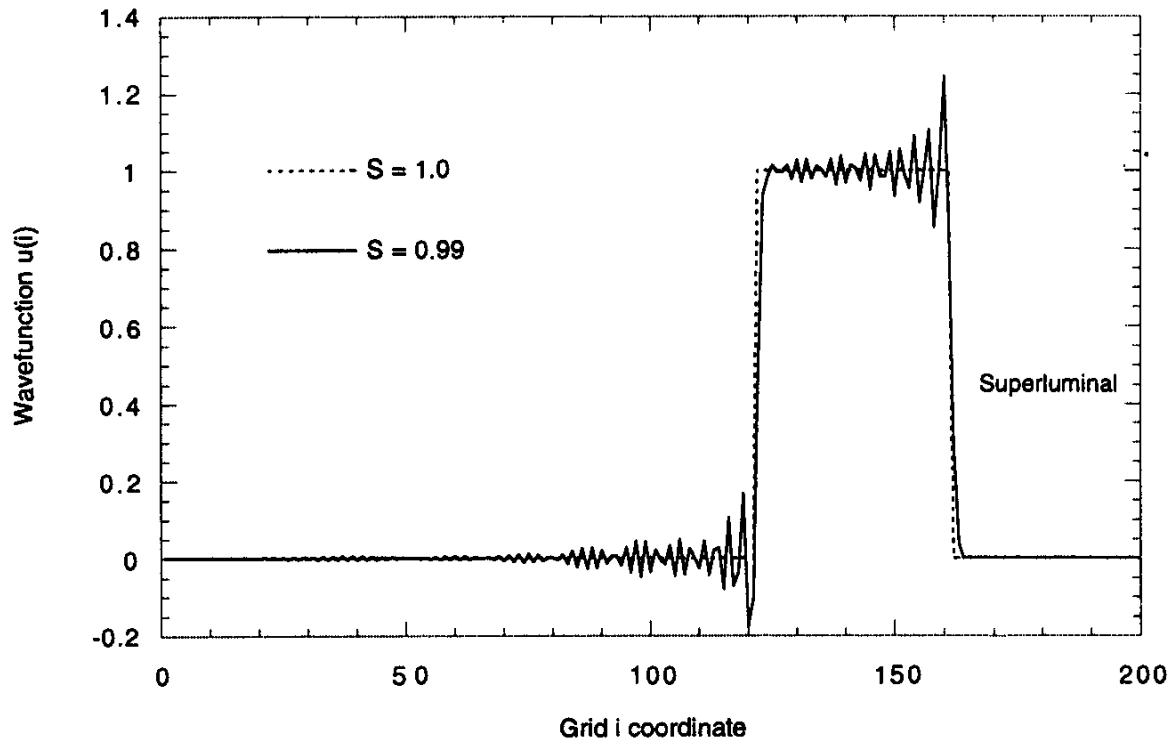
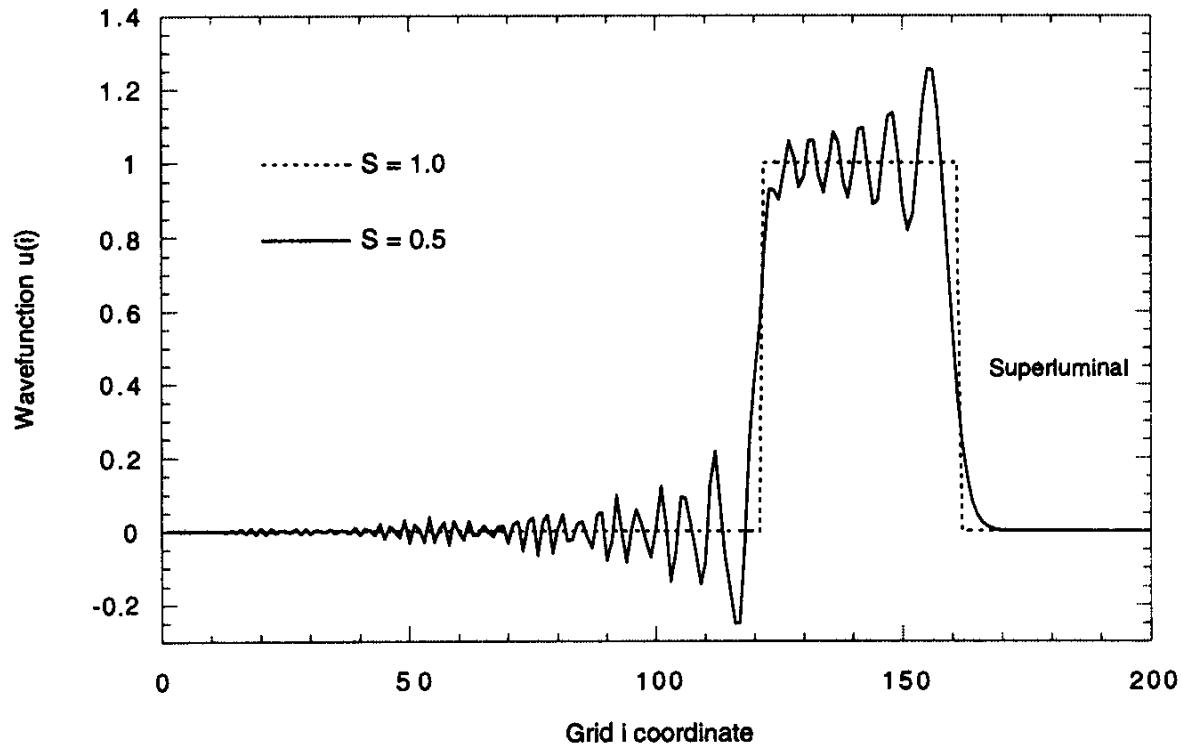
(a) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.99$.(b) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.5$.

Fig. 2.3 Effect of numerical dispersion upon a rectangular pulse propagating in free space for three different Courant stability factors: $S = 1$ (Δt equal to the magic time-step), $S = 0.99$ (Δt just 1% below the magic time-step), and $S = 0.5$ (Δt 50% below the magic time-step).

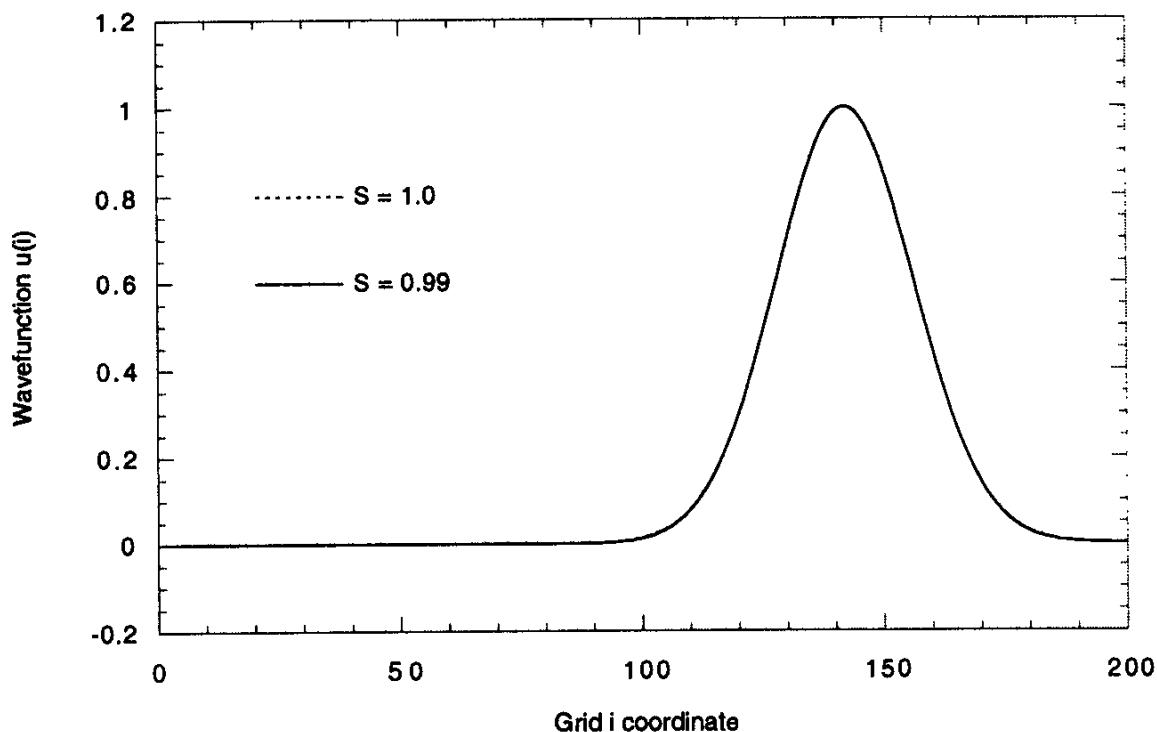
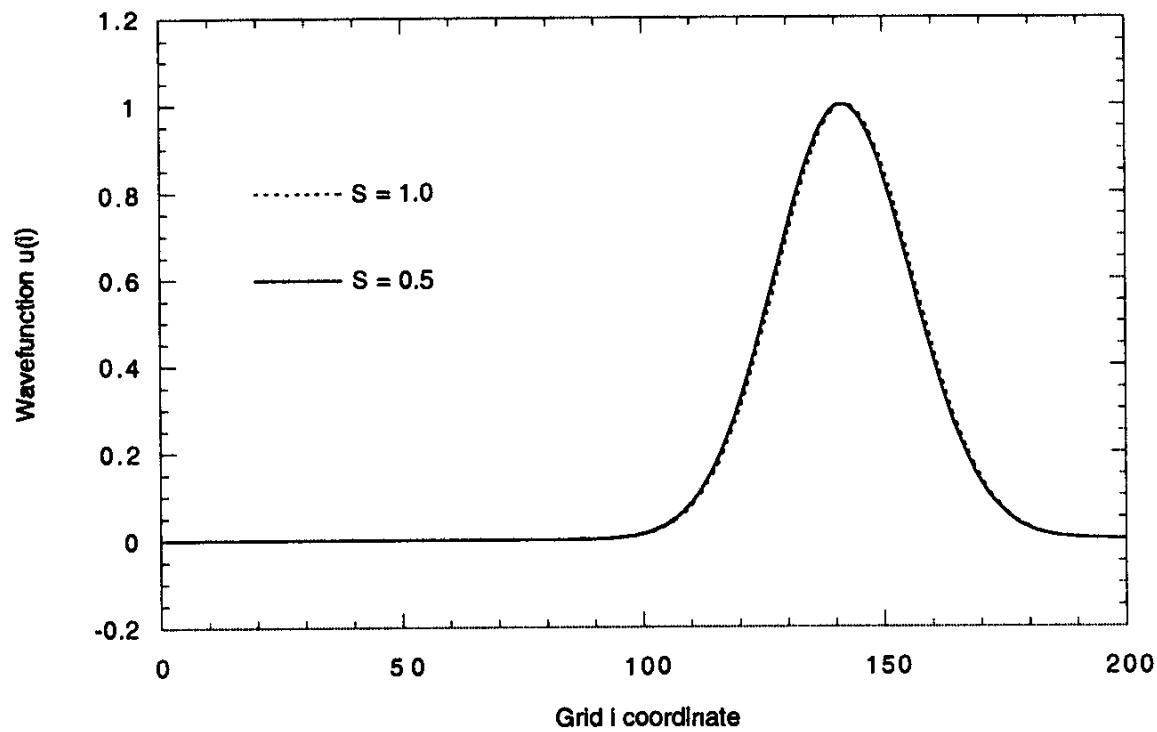
(a) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.99$.(b) Comparison of calculated pulse propagation for $S = 1$ and $S = 0.5$.

Fig. 2.4 Effect of numerical dispersion upon a Gaussian pulse propagating in free space for three different Courant stability factors: $S = 1$ (Δt equal to the magic time-step), $S = 0.99$ (Δt just 1% below the magic time-step), and $S = 0.5$ (Δt 50% below the magic time-step).

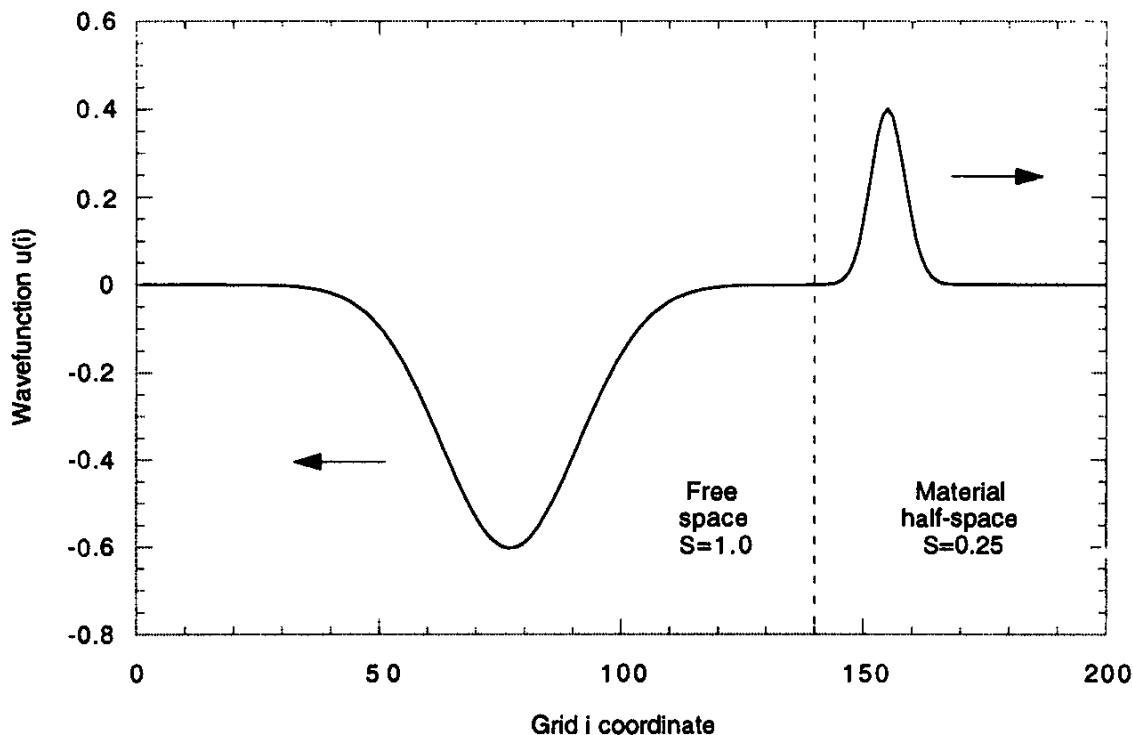


Fig. 2.5 Example of the calculated reflection and transmission of a Gaussian pulse at an interface between free space and a lossless material half-space having $v_p = c/4$.

Fig. 2.5 graphs the calculated transmission and reflection of a Gaussian pulse at the interface from free space to a lossless material half-space having a phase velocity $v_p = c/4$. Here, the incident pulse has a unity amplitude and a spatial width of $40\Delta x$ between its $1/e$ points. Both regions of the grid are modeled using a uniform Δx and Δt . The physics of the phase-velocity discontinuity between the two regions is modeled simply by assigning a different Courant number $S = c\Delta t/\Delta x$ to grid-points in each region when implementing the time-stepping algorithm of (2.16). In the free-space region (grid-points between $i = 1$ and $i = 139$) we assume $S = 1$, whereas in the material region (grid-points between $i = 140$ and $i = 200$) we assume $S = 0.25$.

At the time of the snapshot of Fig. 2.5, the incident pulse has already reached the interface at $i = 140$ and experienced partial reflection and transmission. With the material half-space being lossless and nondispersive, the reflected and transmitted pulses must also have a Gaussian shape. Simply by taking the ratio of the peak values of the reflected and incident pulses, a reflection coefficient of -0.603 is obtained. This is within 0.5% of the exact value of -0.6 . Similarly, taking the ratio of the peak values of the transmitted and incident pulses leads to a transmission coefficient within 0.5% of the exact value of 0.4 . Finally, the spatial width of the reflected pulse is virtually identical to that of the incident, while the spatial width of the transmitted pulse is almost exactly 0.25 times that of the incident, as predicted by the analytical theory. This simple exercise is our first indication that a numerical model of the wave equation can simulate the scattering properties of material structures by assigning appropriate updating coefficients to the time-stepping algorithm of each grid-point.

2.7 NUMERICAL STABILITY

We have seen that the choice of space increment Δx and time-step Δt can affect the velocity of propagation of numerical waves in a finite-difference grid modeling the one-dimensional scalar wave equation (2.1), and therefore the numerical error. In this section, we show that another consideration enters into the selection of the time-step: Δt must be bounded in order to ensure numerical stability. The following definition [3] forms the basis of this discussion:

An explicit numerical solution of (2.1) such as (2.16) is *stable* if it produces a bounded result given a bounded input. The numerical solution is *unstable* if it produces an unbounded result given a bounded input.

Numerical instability is an undesirable possibility with explicit numerical differential equation solvers that can cause the computed results to spuriously increase without limit as time-marching continues.

A classical approach to analyze numerical stability is the spectral technique developed by von Neumann. This method expresses the error in a numerical solution such as (2.16) at any point in time as a finite spatial Fourier series. Numerical stability results if each Fourier term has a unity-or-less growth factor over one time-step. Then, assuming that each Fourier term is initially bounded, each term remains bounded at all subsequent time-steps. Since the system is linear, the total error represented by the finite sum of the Fourier terms must also be bounded at any time-step. Hence, the numerical solution satisfies the stability definition given above.

In this section, however, we take a different (yet rigorous) approach to numerical stability that is more intuitive to students than von Neumann's technique. The alternative approach avoids the need for an eigenvalue / eigenspectrum analysis and also allows straightforward estimates of the growth rate of unstable solutions. The method is based upon a complex-frequency analysis of numerical dispersion relation (2.22). Its application to Maxwell's equations and the Yee algorithm will be discussed in detail in Chapter 4.

2.7.1 Complex-Frequency Analysis

We again consider the sinusoidal traveling wave of (2.6) present in our finite-difference grid and discretely sampled at (x_i, t_n) . Now, however, we allow for the possibility of a complex-valued numerical angular frequency, $\tilde{\omega} = \tilde{\omega}_{real} + j\tilde{\omega}_{imag}$. Then, (2.20) becomes

$$\begin{aligned} u_i^n &= e^{j(\tilde{\omega} n \Delta t - \tilde{k} i \Delta x)} \\ &= e^{j[(\tilde{\omega}_{real} + j\tilde{\omega}_{imag})n \Delta t - \tilde{k} i \Delta x]} = e^{-\tilde{\omega}_{imag} n \Delta t} e^{j(\tilde{\omega}_{real} n \Delta t - \tilde{k} i \Delta x)} \end{aligned} \quad (2.40)$$

Here, as in our earlier analysis, \tilde{k} is the wavenumber of the numerical sinusoidal traveling wave. We note that (2.40) permits either a constant wave amplitude with time ($\tilde{\omega}_{imag} = 0$), an exponentially decreasing amplitude with time ($\tilde{\omega}_{imag} > 0$), or an exponentially increasing amplitude with time ($\tilde{\omega}_{imag} < 0$).

Given this basis, we proceed to analyze numerical dispersion relation (2.22) allowing for a complex-valued angular frequency:

$$\cos(\tilde{\omega}\Delta t) = \left(\frac{c\Delta t}{\Delta x}\right)^2 [\cos(\tilde{k}\Delta x) - 1] + 1 \quad (2.41)$$

Again defining the Courant stability factor $S = c\Delta t/\Delta x$, we solve (2.41) for $\tilde{\omega}$:

$$\begin{aligned} \tilde{\omega} &= \frac{1}{\Delta t} \cos^{-1} \left\{ S^2 [\cos(\tilde{k}\Delta x) - 1] + 1 \right\} \\ &= \frac{1}{\Delta t} \cos^{-1}(\xi) = \frac{1}{\Delta t} \left[\frac{\pi}{2} - \sin^{-1}(\xi) \right] \end{aligned} \quad (2.42a)$$

where

$$\xi = S^2 [\cos(\tilde{k}\Delta x) - 1] + 1 \quad (2.42b)$$

We observe from (2.42b) that $1 - 2S^2 \leq \xi \leq 1$ for all possible real values of \tilde{k} , that is, those numerical waves having zero exponential attenuation per grid space cell. (For this case, the cosine function is bounded by $-1 \leq \cos(k\Delta x) \leq 1$.) Now, consider dividing the $1 - 2S^2 \leq \xi \leq 1$ range into two subranges, as follows:

- (a) $-1 \leq \xi \leq 1$. This subrange exists for $0 \leq S \leq 1$. Here, $\sin^{-1}(\xi)$ is real-valued and hence, real values of $\tilde{\omega}$ are obtained in (2.42a). With $\tilde{\omega}_{imag} = 0$, Equation (2.40) yields a constant wave amplitude with time.
- (b) $1 - 2S^2 \leq \xi < -1$. This subrange exists for $1 - 2S^2 < -1$, or equivalently $S > 1$. The most negative value in this subrange, $\xi_{lower bound}$, occurs for $\cos(\tilde{k}\Delta x) = -1$, i.e., $\tilde{k}\Delta x = \pi$ (equivalent to a numerical wavelength $\tilde{\lambda} = 2\Delta x$.) This lower bound is given by

$$\xi_{lower bound} = 1 - 2S^2 \quad \} \text{ for } \tilde{k}\Delta x = \pi \quad (2.43)$$

Here, $\sin^{-1}(\xi)$ is a complex-valued function given by (2.34), which is repeated here for convenience:

$$\sin^{-1}(\xi) = -j \ln \left(j\xi + \sqrt{1 - \xi^2} \right) \quad (2.44)$$

We sense that Case (b), that is, $S > 1$, may provide unusual numerical wave propagation phenomena. To explore these phenomena, we substitute (2.44) into (2.42a) and solve for $\tilde{\omega}$. This yields

$$\begin{aligned}\tilde{\omega} &= \frac{1}{\Delta t} \left[\frac{\pi}{2} + j \ln \left(j\xi + \sqrt{1 - \xi^2} \right) \right] \\ &= \frac{1}{\Delta t} \left[\frac{\pi}{2} + j \ln \left(j\xi + j\sqrt{\xi^2 - 1} \right) \right]\end{aligned}\quad (2.45a)$$

Factoring out $j = e^{+j\pi/2}$ in the argument of the natural logarithm, we obtain

$$\begin{aligned}\tilde{\omega} &= \frac{1}{\Delta t} \left\{ \frac{\pi}{2} + j \ln \left[\left(\xi + \sqrt{\xi^2 - 1} \right) e^{+j\pi/2} \right] \right\} \\ &= \frac{1}{\Delta t} \left\{ \frac{\pi}{2} + j \ln \left[\left(-\xi - \sqrt{\xi^2 - 1} \right) e^{-j\pi/2} \right] \right\}\end{aligned}\quad (2.45b)$$

Upon taking the natural logarithm, we further obtain

$$\begin{aligned}\tilde{\omega} &= \frac{1}{\Delta t} \left\{ \frac{\pi}{2} + j \left[\ln \left(-\xi - \sqrt{\xi^2 - 1} \right) - j \frac{\pi}{2} \right] \right\} \\ &= \frac{1}{\Delta t} \left\{ \pi + j \ln \left(-\xi - \sqrt{\xi^2 - 1} \right) \right\}\end{aligned}\quad (2.45c)$$

Now, the real and imaginary parts of $\tilde{\omega}$ can be separated:

$$\tilde{\omega}_{real} = \frac{\pi}{\Delta t} ; \quad \tilde{\omega}_{imag} = \frac{1}{\Delta t} \ln \left(-\xi - \sqrt{\xi^2 - 1} \right) \quad (2.46)$$

Finally, substituting (2.46) into (2.40), we obtain

$$\begin{aligned}u_i^n &= e^{\left[-n \ln \left(-\xi - \sqrt{\xi^2 - 1} \right) \right]} e^{j[(\pi/\Delta t)(n\Delta t) - \tilde{k} i\Delta x]} \\ &= \left(\frac{1}{-\xi - \sqrt{\xi^2 - 1}} \right)^{**n} e^{j[(\pi/\Delta t)(n\Delta t) - \tilde{k} i\Delta x]}\end{aligned}\quad (2.47)$$

where $**n$ denotes the n 'th power. From (2.47), we define the following multiplicative factor greater than 1 that amplifies the numerical wave every time step:

$$q_{\text{growth}} \equiv \frac{1}{-\xi - \sqrt{\xi^2 - 1}} = -\xi + \sqrt{\xi^2 - 1} \quad (2.48)$$

Our suspicion that Case (b) is very interesting is now borne out. Because $\xi < -1$ for this case, (2.47) and (2.48) define an *exponential growth* of the numerical wave with time-step number n . We see that the dominant exponential growth occurs for the most negative possible value of ξ , i.e. $\xi_{\text{lower bound}}$ defined in (2.43). As stated earlier in the context of (2.43), this condition occurs for $\tilde{k} = \pi/\Delta x$ (equivalent to a numerical wavelength $\tilde{\lambda} = 2\Delta x$). Substituting $\xi_{\text{lower bound}}$ of (2.43) into (2.48) yields the maximum growth-factor per time-step, valid for the regime $S \geq 1$:

$$\begin{aligned} q_{\text{growth}} &= -\left(1 - 2S^2\right) + \sqrt{\left(1 - 2S^2\right)^2 - 1} \\ &= 2S^2 + 2S\sqrt{S^2 - 1} - 1 \\ &= \left(S + \sqrt{S^2 - 1}\right)^2 \end{aligned} \quad (2.49)$$

From (2.49), we see that a Courant stability factor $S = 1$ yields no solution growth. However, a stability factor only 0.05% larger ($S = 1.0005$) yields a growth-factor of:

- 1.0653 every time-step
- 1.8822 every 10 time-steps
- 558.7 every 100 time-steps
- 2.96×10^{27} every 1,000 time-steps

It is now quite clear that, for *any* Courant factor $S > 1$, our theory predicts an exponentially growing, rapidly oscillating ($\tilde{\lambda} = 2\Delta x$), sinusoidal numerical wave propagating within the computation grid. This is the origin of numerical instability.

A second implication of this theory is that such an exponentially increasing sinusoidal numerical wave has a fixed temporal frequency f_0 that is *independent* of every algorithm parameter but Δt :

$$f_0 = \frac{\tilde{\omega}_{\text{real}}}{2\pi} = \frac{1}{2\Delta t} \quad (2.50)$$

Finally, a third implication of this theory is that the numerical propagation velocity of the exponentially growing wave is given by

$$\tilde{v}_p = \frac{\tilde{\omega}_{\text{real}}}{\tilde{k}} = \frac{(\pi / \Delta t)}{(\pi / \Delta x)} = \frac{\Delta x}{\Delta t} = \frac{c}{S} \quad (2.51)$$

In summary, for computational stability in modeling the one-dimensional scalar wave equation using the numerical algorithm of (2.20), there exists an upper bound on Δt relative to Δx and the free-space speed of light. This bound is given by

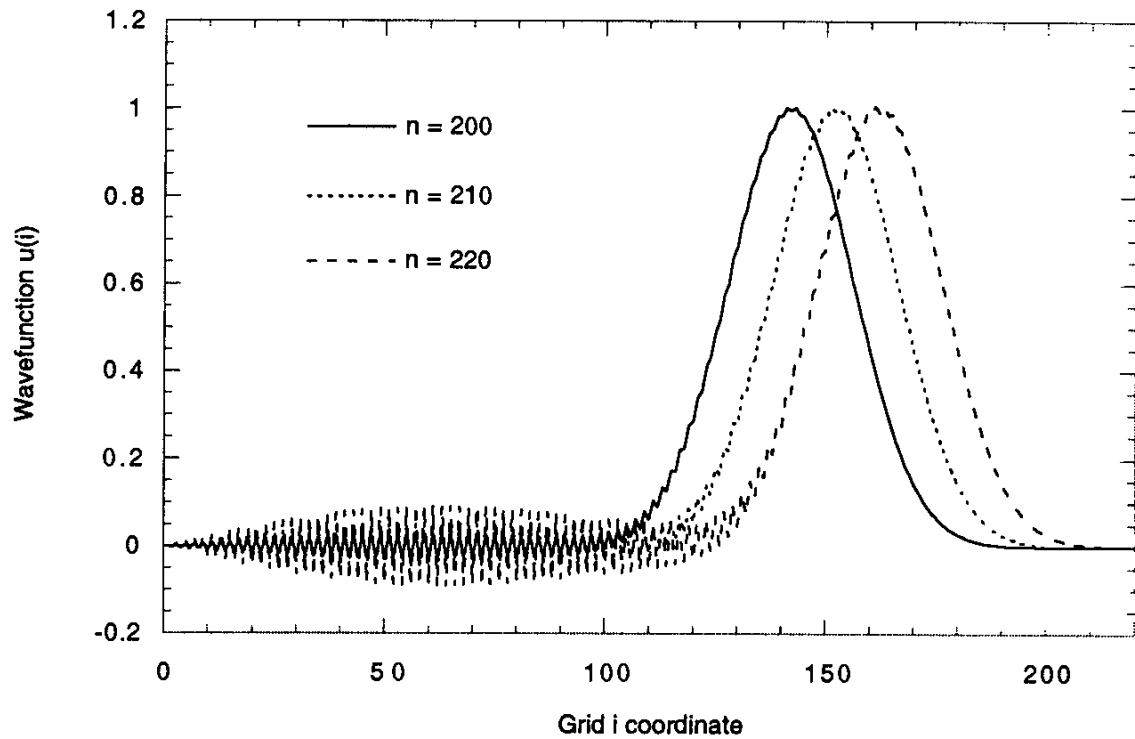
$$S \equiv \frac{c \Delta t}{\Delta x} \leq 1 \quad \longrightarrow \quad \Delta t \leq \frac{\Delta x}{c} \quad (2.52)$$

following the notation of (2.28a) and [1]. S is called the numerical stability factor or Courant number. In most wave-interaction modeling problems, Δx is the first algorithm parameter to be specified by the analyst. Namely, Δx is selected to adequately resolve a structure's geometrical details and/or the principal components of the spectrum of wavelengths propagating within the grid. Once Δx is selected in this manner, the maximum value of Δt is given by (2.52). If Δt is somehow taken to be larger than this bound, then it is a *certainty* that the numerical algorithm will eventually undergo a progressive instability wherein the computed values grow exponentially. Interestingly, the upper bound for stable operation of the algorithm is exactly the magic time-step discussed earlier in this chapter.

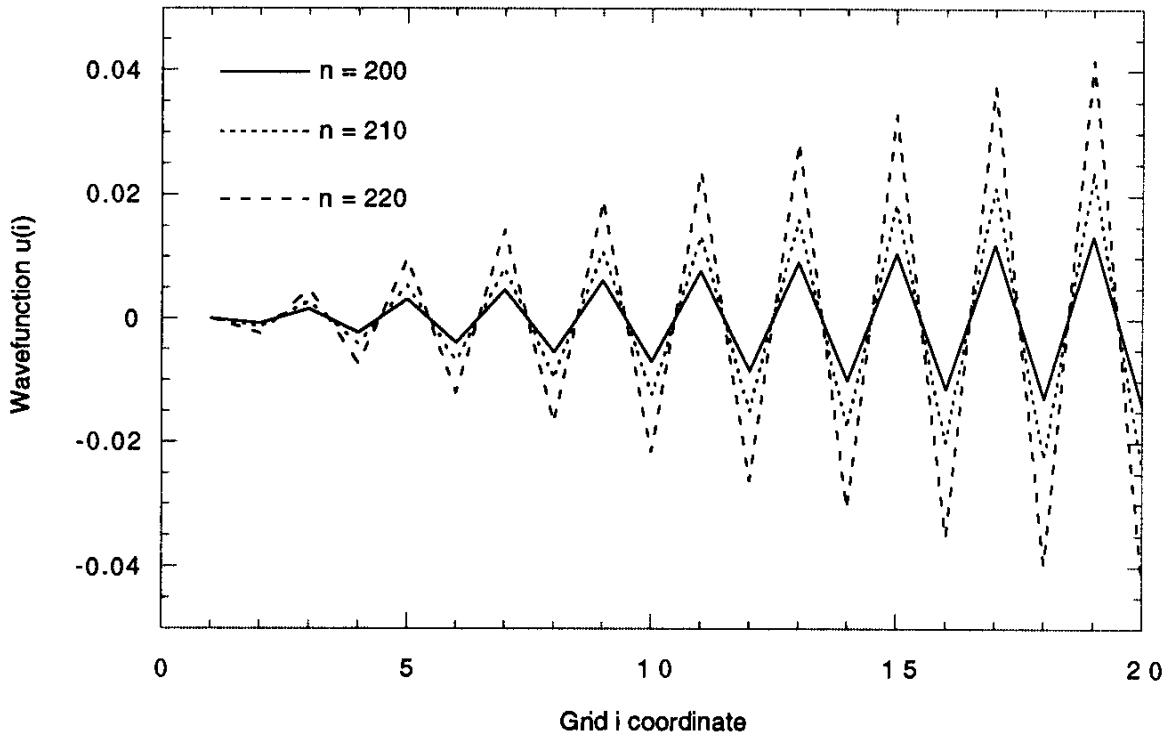
2.7.2 Examples of Calculations Involving Numerical Instability

We first consider an example of the beginning of a numerical instability arising because the Courant stability condition is violated equally at *every* grid-point. Fig. 2.6(a) graphs three snapshots of the free-space propagation of a Gaussian pulse within a uniform grid having the Courant stability factor $S = 1.0005$. The exciting pulse waveform has a $40\Delta t$ temporal width between its $1/e$ points, and reaches its peak value of 1.0 at time-step $n = 60$. Graphs of the wavefunction $u(i)$ versus the grid coordinate i are shown at time-steps $n = 200$, $n = 210$, and $n = 220$. We see that the trailing edge of the Gaussian pulse is contaminated by a rapidly oscillating and growing noise component that does not exist in Fig. 2.4(a), which shows the same Gaussian pulse at the same time but with $S \leq 1.0$. In fact, the noise component in Fig. 2.6(a) results from the onset of numerical instability within the grid due to $S = 1.0005 > 1.0$. Because this noise grows exponentially with time-step number n , it quickly overwhelms the desired numerical results for the propagating Gaussian pulse. Shortly thereafter, the exponential growth of the noise increases the calculated field values beyond the dynamic range of the computer being used, resulting in run-time floating-point overflows and errors.

Fig. 2.6(b) is an expanded view of Fig. 2.6(a) between grid-points $i = 1$ and $i = 20$, showing a segment of the numerical noise on the trailing edge of the Gaussian pulse. We see that the noise oscillates with a spatial period of 2 grid cells, i.e. $\tilde{\lambda} = 2\Delta x$, in accordance with the theory developed in Section 2.7.1. In addition, upon analyzing the raw data underlying Fig. 2.6(b), it is observed that the exponential growth factor q is in the range 1.75–2.0 for every 10-time-step advance of the algorithm. This compares favorably with the theoretical value of 1.88 determined using (2.49).



(a) Comparison of calculated pulse propagation at $n = 200$, 210 , and 220 time-steps over grid coordinates $i = 1$ through $i = 220$.



(b) Expanded view of (a) over grid coordinates $i = 1$ through $i = 20$.

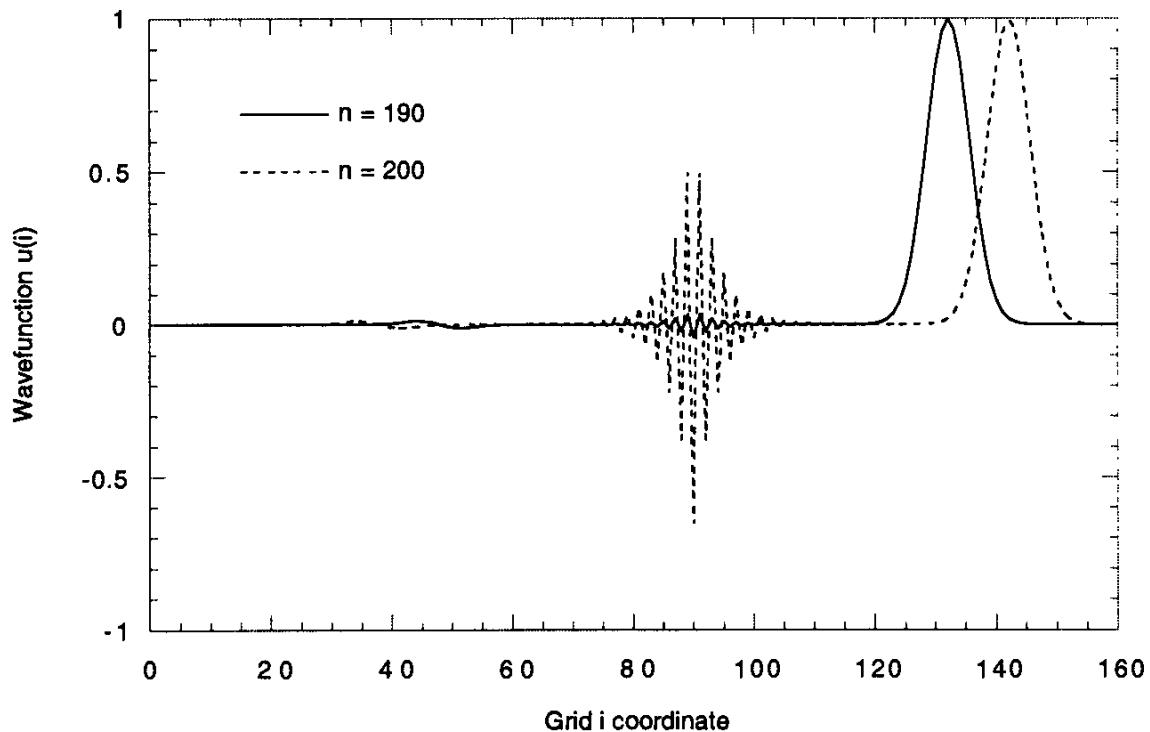
Fig. 2.6 The beginning of numerical instability for a Gaussian pulse propagating in free space. The Courant stability factor is $S = 1.0005$ at each grid-point.

We next consider an example of the beginning of a numerical instability arising because the Courant stability condition is violated at only a *single* grid-point. Fig. 2.7(a) graphs two snapshots of the free-space propagation of a narrow Gaussian pulse within a grid having the Courant factor $S = 1.0$ at all points except at $i = 90$, where $S = 1.075$. The exciting pulse waveform has a $10\Delta t$ temporal width between its $1/e$ points, and reaches its peak value of 1.0 at time-step $n = 60$. Graphs of the wavefunction $u(i)$ versus the grid coordinate i are shown at time-steps $n = 190$ and $n = 200$. In contrast to Fig. 2.6(a), the rapidly oscillating and growing noise component due to numerical instability originates at just a single grid-point along the trailing edge of the Gaussian pulse ($i = 90$) where S exceeds 1.0, rather than along the entirety of the trailing edge. Despite this localization of the source of the instability, the noise again grows exponentially with time-step number n . In this case, the noise propagates symmetrically in both directions from the unstable point. Ultimately, the noise again fills the entire grid, overwhelms the desired numerical results for the propagating Gaussian pulse, and causes run-time floating-point overflows.

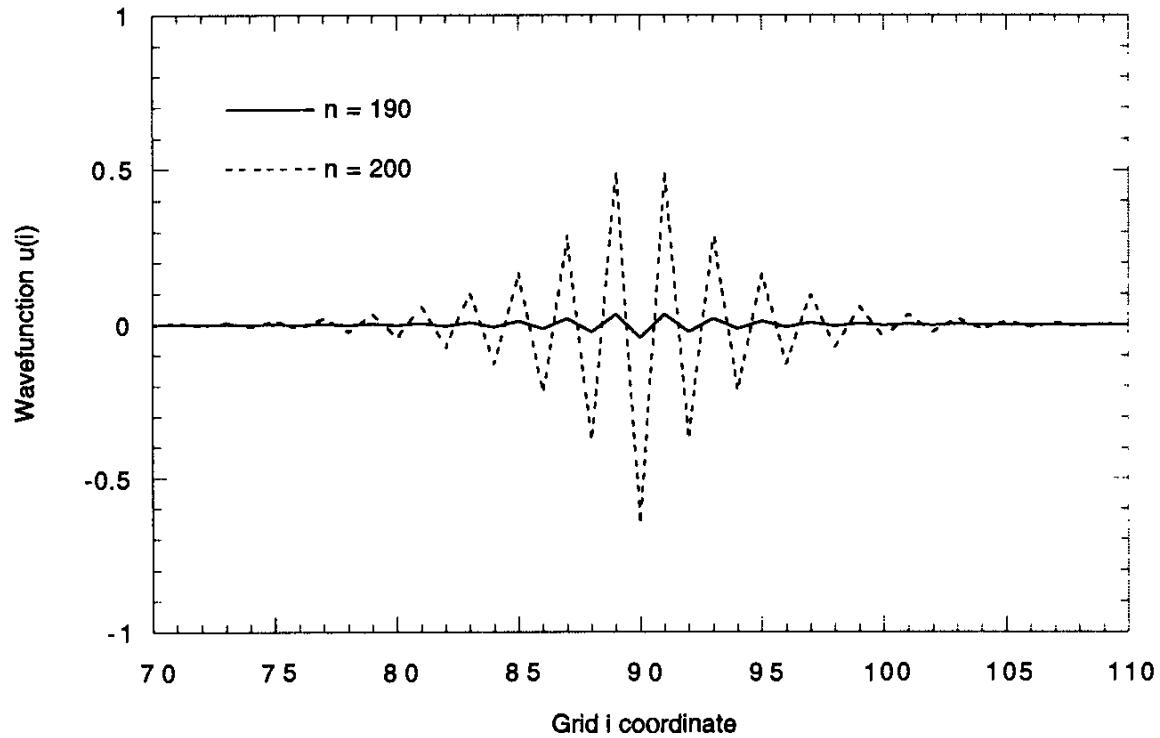
Fig. 2.7(b) is an expanded view of Fig. 2.7(a) between grid-points $i = 70$ and $i = 110$, showing how the calculated noise due to the numerical instability originates at grid-point $i = 90$. Again, the noise oscillates with a spatial period of 2 grid cells, i.e. $\tilde{\lambda} = 2\Delta x$. However, the rate of exponential growth here is much less than that predicted by the theory of Section 2.7.1, wherein *all* grid-points were assumed to violate Courant stability. Upon analyzing the raw data underlying Fig. 2.7(b), the exponential growth factor $q \cong 15.0$ is observed for every 10-time-step advance of the algorithm. This compares to $q \cong 2,205$ per 10 time-steps determined by substituting $S = 1.075$ into (2.49). Thus, it is clear that a grid having one or just a few localized points of numerical instability can “blow up” much more slowly than a uniformly unstable grid having a comparable or even smaller Courant factor S .

2.8 SUMMARY

This chapter considered the numerical finite-difference time-domain (FDTD) solution of the one-dimensional scalar wave equation. First, the analytical propagating-wave solutions were obtained. Then, finite differences were introduced and applied to the wave equation, leading to introductory discussions of numerical dispersion, numerical phase velocity, the “magic” time-step, and numerical stability. Calculated examples of impulsive wave propagation were provided to illustrate these fundamental concepts. The discussions of this chapter serve as the fundamental basis for our work in the following chapters, with similar concepts regarding FDTD computational solution of the vector Maxwell’s equations in two and three dimensions.



(a) Comparison of calculated pulse propagation at $n = 190$ and $n = 200$ time-steps over grid coordinates $i = 1$ through $i = 220$.



(b) Expanded view of (a) over grid coordinates $i = 70$ through $i = 110$.

Fig. 2.7 The beginning of numerical instability for a Gaussian pulse propagating in free space. Unlike Fig. 2.6, the Courant stability factor is $S = 1$ at all grid-points but $i = 90$, where $S = 1.075$.

APPENDIX 2A: ORDER OF ACCURACY¹

This appendix qualitatively defines what we mean by the “order of accuracy” of a numerical approximation of a wavelike equation. This discussion is abstracted from Sections 2.1 and 2.2 of [4], which provide considerable supporting mathematical detail.

We consider a general first-order system of equations for the wavefunction u , which we symbolically denote by

$$\frac{\partial u}{\partial t} = \mathcal{L}u \quad (2A.1)$$

where \mathcal{L} is a general linear operator, for example, the spatial-derivative portion of Maxwell’s equations. We have a one-step approximation given by

$$\frac{u^{n+1} - u^n}{\Delta t} = \mathcal{L}_h u \quad (2A.2)$$

Now, assume that u is a sufficiently smooth solution of (2A.1). Let Δt denote the time step and h the mesh size in all spatial directions. Then, the approximation \mathcal{L}_h has order of accuracy p in time and q in space if the truncation error τ satisfies

$$\tau = \frac{u^{n+1} - u^n}{\Delta t} - \mathcal{L}_h u = O[(\Delta t)^p + h^q] \quad (2A.3)$$

We usually calculate the order of accuracy (p, q) by a Taylor series. Note that if the solution is not sufficiently smooth, then the order of accuracy is reduced. This can happen if the coefficients are not smooth, for example, due to the permittivity jump across the interface of two dielectric media.

2A.1 Lax-Richtmyer Equivalence Theorem

The local definition of accuracy is important because of the Lax-Richtmyer equivalence theorem [5, 6]. This theorem states that if a scheme has a truncation error of order (p, q) and the scheme is stable, then the difference between the analytic solution and the numerical solution in an appropriate norm is of the order $(\Delta t)^p + h^q$ for all finite time. This equivalence theorem can be extended to variable coefficients and even smooth nonlinear problems. Further, Gustafsson has shown that if the numerical boundary treatment is one order less accurate than the interior accuracy, then the order of the global accuracy is preserved [7].

¹This appendix is contributed by Eli Turkel.

We note that the Lax-Richtmyer equivalence theorem deals with errors measured in some norm, for example, least squares. Hence, we deal with the total error. At times, we may be mainly interested in either the phase (velocity) of the solution or its amplitude rather than the entire information. In fact, it is possible to construct schemes that have higher-order accuracy in their phase error than when measured in the least-squares norm [4].

For numerical schemes approximating hyperbolic, wavelike equations, the order of accuracy in time and space are often taken to be equal, for example, (2, 2). This is because hyperbolic systems behave in a similar manner in both time and space. In fact for the simplest equation, $\partial u / \partial t = \partial u / \partial x$, one cannot distinguish between the time and space directions. Nevertheless, it is often useful to consider schemes that have higher order accuracy in space than in time, for example, (2, 4). This is because improving the order of temporal accuracy decreases the work load only in the time direction without affecting the field storage. On the other hand, improving the order of spatial accuracy enables the use of a coarser mesh, which decreases the work in each space dimension and also decreases the field storage. Hence, numerical accuracy in space is more important than numerical accuracy in time.

2A.2 Limitations

Since the order of accuracy is found by a Taylor series expansion, this “figure of merit” only gives useful information in the limit of sufficiently fine meshes. Equivalently, it stresses the lower spatial frequencies. For high spatial frequencies not resolved by the mesh, we note that the order of accuracy has no meaning. In fact, one can construct numerical methods that are formally low-order accurate but have better accuracy for the higher spatial frequencies [8, 9]. For such approaches, there may be no equivalent of the Lax-Richtmyer theorem.

REFERENCES

- [1] Schneider, J. B., and C. L. Wagner, “FDTD dispersion revisited: Faster-than-light propagation,” *IEEE Microwave and Guided Wave Lett.*, Vol. 9, 1999, pp. 54–56.
- [2] Churchill, R. V., J. W. Brown, and R. F. Verhey, *Complex Variables and Applications*. New York: McGraw-Hill, 1976.
- [3] Hoffman, J., *Numerical Methods for Engineers and Scientists*. New York: McGraw-Hill, 1992.
- [4] Turkel, E., “High-Order Methods,” Chap. 2 in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, A. Taflove, ed., Norwood, MA: Artech House, 1998.
- [5] Lax, P. D., and L. Nirenberg, “On stability of difference schemes: A sharp form of Garding’s inequality,” *Commun. on Pure and Applied Mathematics*, Vol. 19, 1966, pp. 473–492.
- [6] Richtmyer, R. D., and K. W. Morton, *Difference Methods for Initial-Value Problems*, 2nd ed., New York: Wiley-Interscience, 1967.
- [7] Gustafsson, B., H.-O. Kreiss, and J. Oliger, *Time-Dependent Problems and Difference Methods*, New York: Wiley-Interscience, 1995.

- [8] Tam, C. K. W., and J. C. Webb, "Dispersion-relation preserving finite-difference schemes for computational acoustics," *J. Computational Physics*, Vol. 107, 1993, pp. 262–281.
- [9] Zingg, D. W., "A review of high-order and optimized finite-difference methods for simulating linear wave phenomena," AIAA paper 97-2088, 1997.

BIBLIOGRAPHY ON STABILITY OF FINITE-DIFFERENCE METHODS

- Carnahan, B., H. A. Luther, and J. O. Wilkes, *Applied Numerical Methods*, New York: Wiley, 1969.
- Douglas, J., Jr., "On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations," *J. Soc. Indust. Appl. Math.*, Vol. 4, 1956, pp. 20–37.
- Forsythe, G. E., and W. R. Wasow, *Finite-Difference Methods for Partial Differential Equations*, New York: Wiley, 1960.
- Gustafsson, B., H. O. Kreiss, and A. Sundstrom, "Stability theory of difference approximations for mixed initial-boundary value problems, II," *Mathematics of Computation*, Vol. 26, 1972, pp. 649–686.
- O'Brien, G., M. Hyman, and S. Kaplan, "A study of the numerical solution of partial differential equations," *J. Math. Phys.*, Vol. 29, 1951, pp. 233–251.
- Peaceman, D. W., and H. H. Rachford, Jr., "The numerical solution of parabolic and elliptic differential equations," *J. Soc. Indust. Appl. Math.*, Vol. 3, 1955, pp. 28–41.
- Richtmyer, R. D., *Difference Methods for Initial Value Problems*, New York: Wiley-Interscience, 1957.
- Strikwerda, J. C., *Finite Difference Schemes and Partial Differential Equations*, New York: Wadsworth & Brooks/Cole Mathematics Series (ISBN 0-534-09984-X), 1989.

PROBLEMS

- 2.1 Use the Taylor's series expansion method to derive a second-order accurate central-difference approximation for $\partial u / \partial x$.
- 2.2 Repeat Problem 2.1, but now derive a fourth-order accurate central-difference approximation for $\partial u / \partial x$.
- 2.3 Replicate the graphical results of Fig. 2.1.
- 2.4 Develop graphical results similar to those of Fig. 2.1, but for a Courant stability factor $S = 1/\sqrt{2}$.
- 2.5 Replicate the graphical results of Fig. 2.2.
- 2.6 Develop graphical results similar to those of Fig. 2.2, but for a Courant stability factor $S = 1/\sqrt{2}$.
- 2.7 Write a computer program that implements the solution of the one-dimensional scalar wave equation discussed in this chapter. To source a wave, specify u_0^n at the left boundary of the grid. Test your program by replicating the graphical results of Figs. 2.3(a) and 2.3(b).

- 2.8 Use the scalar-wave-equation computer program developed in Problem 2.7 to replicate the graphical results of Fig. 2.4(a) and 2.4(b).
- 2.9 Use the scalar-wave-equation computer program developed in Problem 2.7 to replicate the graphical results of Fig. 2.5.
- 2.10 Use the scalar-wave-equation computer program developed in Problem 2.7 to replicate the graphical results of Fig. 2.6(a) and 2.6(b).
- 2.11 Use the scalar-wave-equation computer program developed in Problem 2.7 to replicate the graphical results of Fig. 2.7(a) and 2.7(b).

Chapter 3

Introduction to Maxwell's Equations and the Yee Algorithm

3.1 INTRODUCTION

In this chapter, we consider the foundation of FDTD electromagnetic field analysis, the algorithm introduced by Kane Yee in 1966 [1]. Yee's insight was to choose a geometric relation for his spatial sampling of the vector components of the electric and magnetic fields that robustly represents both the differential and integral forms of Maxwell's equations. Many alternative griddings of Maxwell's equations have been proposed in the approximately 35 years since his paper. However, none have had the seminal impact and longevity of the "original family recipe."

3.2 MAXWELL'S EQUATIONS IN THREE DIMENSIONS

Consider a region of space that has no electric or magnetic current sources, but may have materials that absorb electric or magnetic field energy. Then using MKS units, the time-dependent Maxwell's equations are given in differential and integral form by

Faraday's Law:

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} - \vec{M} \quad (3.1a)$$

$$\frac{\partial}{\partial t} \iint_A \vec{B} \cdot d\vec{A} = - \oint_{\ell} \vec{E} \cdot d\vec{\ell} - \iint_A \vec{M} \cdot d\vec{A} \quad (3.1b)$$

Ampere's Law:

$$\frac{\partial \vec{D}}{\partial t} = \nabla \times \vec{H} - \vec{J} \quad (3.2a)$$

$$\frac{\partial}{\partial t} \iint_A \vec{D} \cdot d\vec{A} = \oint_{\ell} \vec{H} \cdot d\vec{l} - \iint_A \vec{J} \cdot d\vec{A} \quad (3.2b)$$

Gauss' Law for the electric field:

$$\nabla \cdot \vec{D} = 0 \quad (3.3a)$$

$$\iint_A \vec{D} \cdot d\vec{A} = 0 \quad (3.3b)$$

Gauss' Law for the magnetic field:

$$\nabla \cdot \vec{B} = 0 \quad (3.4a)$$

$$\iint_A \vec{B} \cdot d\vec{A} = 0 \quad (3.4b)$$

In (3.1)–(3.4), the following symbols (and their MKS units) are defined:

- \vec{E} : electric field (volts / meter)
- \vec{D} : electric flux density (coulombs / meter²)
- \vec{H} : magnetic field (amperes / meter)
- \vec{B} : magnetic flux density (webers / meter²)
- A : arbitrary three-dimensional surface
- $d\vec{A}$: differential normal vector that characterizes surface A (meter²)
- ℓ : closed contour that bounds surface A
- $d\vec{l}$: differential length vector that characterizes contour ℓ (meters)
- \vec{J} : electric current density (amperes / meter²)
- \vec{M} : equivalent magnetic current density (volts / meter²)

In linear, isotropic, nondispersive materials (i.e., materials having field-independent, direction-independent, and frequency-independent electric and magnetic properties), we can relate \vec{D} to \vec{E} and \vec{B} to \vec{H} using simple proportions:

$$\vec{D} = \epsilon \vec{E} = \epsilon_r \epsilon_0 \vec{E} ; \quad \vec{B} = \mu \vec{H} = \mu_r \mu_0 \vec{H} \quad (3.5)$$

where

- ϵ : electrical permittivity (farads / meter)
- ϵ_r : relative permittivity (dimensionless scalar)
- ϵ_0 : free-space permittivity (8.854×10^{-12} farads / meter)
- μ : magnetic permeability (henrys / meter)
- μ_r : relative permeability (dimensionless scalar)
- μ_0 : free-space permeability ($4\pi \times 10^{-7}$ henrys / meter)

Note that \vec{J} and \vec{M} can act as *independent sources* of E - and H -field energy, \vec{J}_{source} and \vec{M}_{source} . We also allow for materials with isotropic, nondispersive electric and magnetic losses that attenuate E - and H -fields via conversion to heat energy. This yields:

$$\vec{J} = \vec{J}_{\text{source}} + \sigma \vec{E} ; \quad \vec{M} = \vec{M}_{\text{source}} + \sigma^* \vec{H} \quad (3.6)$$

where

- σ : electric conductivity (siemens / meter)
- σ^* : equivalent magnetic loss (ohms / meter)

Finally, we substitute (3.5) and (3.6) into (3.1a) and (3.2a). This yields Maxwell's curl equations in linear, isotropic, nondispersive, lossy materials:

$$\frac{\partial \vec{H}}{\partial t} = -\frac{1}{\mu} \nabla \times \vec{E} - \frac{1}{\mu} (\vec{M}_{\text{source}} + \sigma^* \vec{H}) \quad (3.7)$$

$$\frac{\partial \vec{E}}{\partial t} = \frac{1}{\epsilon} \nabla \times \vec{H} - \frac{1}{\epsilon} (\vec{J}_{\text{source}} + \sigma \vec{E}) \quad (3.8)$$

We now write out the vector components of the curl operators of (3.7) and (3.8) in Cartesian coordinates. This yields the following system of six coupled scalar equations:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - (M_{\text{source}_x} + \sigma^* H_x) \right] \quad (3.9a)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - (M_{\text{source}_y} + \sigma^* H_y) \right] \quad (3.9b)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{\text{source}_z} + \sigma^* H_z) \right] \quad (3.9c)$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - (J_{\text{source}_x} + \sigma E_x) \right] \quad (3.10a)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (3.10b)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{\text{source}_z} + \sigma E_z) \right] \quad (3.10c)$$

The system of six coupled partial differential equations of (3.9) and (3.10) forms the basis of the FDTD numerical algorithm for electromagnetic wave interactions with general three-dimensional objects. The FDTD algorithm need not explicitly enforce the Gauss' Law relations indicating zero free electric and magnetic charge, (3.3) and (3.4). This is because these relations are theoretically a direct consequence of the curl equations, as can be readily shown. However, the FDTD space grid must be structured so that the Gauss' Law relations are *implicit* in the positions of the E - and H -field vector components in the grid, and in the numerical space-derivative operations upon these components that model the action of the curl operator. This will be discussed in Section 3.6.9 in the context of the Yee mesh.

Before proceeding with the full three-dimensional FDTD algorithm, it is instructive to consider simplified two-dimensional and one-dimensional cases. These demonstrate important electromagnetic wave phenomena and can yield insight into the analytical and algorithmic features of the general three-dimensional case.

3.3 REDUCTION TO TWO DIMENSIONS

Let us assume that the structure being modeled extends to infinity in the z -direction with no change in the shape or position of its transverse cross section. If the incident wave is also uniform in the z -direction, then all partial derivatives of the fields with respect to z must equal zero. Under these conditions, the full set of Maxwell's curl equations given by (3.9) and (3.10) reduces to

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left[-\frac{\partial E_z}{\partial y} - (M_{\text{source}_x} + \sigma^* H_x) \right] \quad (3.11a)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_z}{\partial x} - (M_{\text{source}_y} + \sigma^* H_y) \right] \quad (3.11b)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{\text{source}_z} + \sigma^* H_z) \right] \quad (3.11c)$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_z}{\partial y} - (J_{\text{source}_x} + \sigma E_x) \right] \quad (3.12a)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon} \left[-\frac{\partial H_z}{\partial x} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (3.12b)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{\text{source}_z} + \sigma E_z) \right] \quad (3.12c)$$

3.3.1 TM_z Mode

Consider grouping the equations of (3.11) and (3.12) according to field vector components. For example, let us first group (3.11a), (3.11b), and (3.12c), which involve only H_x , H_y , and E_z . We shall designate this set of field components the *transverse-magnetic mode with respect to z* (TM_z) in two dimensions:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left[-\frac{\partial E_z}{\partial y} - (M_{\text{source}_x} + \sigma^* H_x) \right] \quad (3.13a)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_z}{\partial x} - (M_{\text{source}_y} + \sigma^* H_y) \right] \quad (3.13b)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - (J_{\text{source}_z} + \sigma E_z) \right] \quad (3.13c)$$

3.3.2 TE_z Mode

Second, let us group (3.12a), (3.12b), and (3.11c) which involve only E_x , E_y , and H_z . We shall designate this set of field components the *transverse-electric mode with respect to z* (TE_z) in two dimensions:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_z}{\partial y} - (J_{\text{source}_x} + \sigma E_x) \right] \quad (3.14a)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon} \left[-\frac{\partial H_z}{\partial x} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (3.14b)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - (M_{\text{source}_z} + \sigma^* H_z) \right] \quad (3.14c)$$

We observe that the TM_z and TE_z modes contain no common field vector components. Thus, these modes can exist simultaneously with *no* mutual interactions for structures composed of isotropic materials or anisotropic materials having no off-diagonal components in the constitutive tensors. The TM_z and TE_z modes constitute the two possible ways that two-dimensional electromagnetic wave interaction problems can be set up for the case of zero partial derivatives in the z -direction.

Physical phenomena associated with these two modes can be very different. This is due to the orientation of the E - and H -field lines relative to the surface of the structure being modeled. We note that the TE_z mode sets up E -field lines in a plane perpendicular to the infinitely long axis (the z -axis) of the structure. If the structure is metallic, a substantial E -field can be supported immediately adjacent and perpendicular to the structure surface without violating the boundary condition of zero E -field tangential to a perfectly conducting surface. As a result, the TE_z mode can support propagating electromagnetic fields bound closely to, or guided by, the surface of a metal structure (the “creeping wave” being a classic example for curved metal surfaces). On the other hand, the TM_z mode sets up E -field lines only parallel to the z -axis. These lines cannot be perpendicular to the structure surface and therefore must be negligible at the surface if it is metallic. This diminishes or eliminates bound or guided near-surface propagating waves for metal surfaces. The presence or absence of surface-type waves can have important implications for scattering and radiation problems.

3.4 REDUCTION TO ONE DIMENSION

3.4.1 x -Directed, z -Polarized TEM Mode

Let us further assume that neither the electromagnetic field excitation nor the modeled geometry has any variation in the y -direction. In effect, we assume that all field partial derivatives with respect to both y and z equal zero, and that the interaction structure consists of an infinite space having possible material layering in the x -direction. Then the two-dimensional TM_z mode of Maxwell's equations in rectangular coordinates given by (3.13) reduces to

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu} \left(M_{\text{source}_x} + \sigma^* H_x \right) \quad (3.15a)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_z}{\partial x} - \left(M_{\text{source}_y} + \sigma^* H_y \right) \right] \quad (3.15b)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_y}{\partial x} - \left(J_{\text{source}_z} + \sigma E_z \right) \right] \quad (3.15c)$$

Assuming that $M_{\text{source}_x} = 0$ for all time and $H_x = 0$ at $t = 0$, then (3.15a) implies that $\partial H_x / \partial t = 0$ at $t = 0$. With no change in H_x at the beginning of the observation, it remains at zero. In fact, using a simple inductive argument, we can show that $H_x = 0$ during the entire observation. Now we have a set of only two equations involving H_y and E_z . We designate the mode determined by this set as an x -directed, z -polarized transverse electromagnetic (TEM) wave in one dimension:

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left[\frac{\partial E_z}{\partial x} - (M_{\text{source}_y} + \sigma^* H_y) \right] \quad (3.16a)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_y}{\partial x} - (J_{\text{source}_z} + \sigma E_z) \right] \quad (3.16b)$$

3.4.2 x -Directed, y -Polarized TEM Mode

Again assuming that all partial derivatives with respect to y equal zero, the two-dimensional TE_z mode of Maxwell's equations in rectangular coordinates given by (3.14) reduces to

$$\frac{\partial E_x}{\partial t} = -\frac{1}{\epsilon} (J_{\text{source}_x} + \sigma E_x) \quad (3.17a)$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon} \left[-\frac{\partial H_z}{\partial x} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (3.17b)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left[-\frac{\partial E_y}{\partial x} - (M_{\text{source}_z} + \sigma^* H_z) \right] \quad (3.17c)$$

Assuming that $J_{\text{source}_x} = 0$ for all time and $E_x = 0$ at $t = 0$, then (3.17a) implies that $\partial E_x / \partial t = 0$ at $t = 0$. Via an inductive argument similar to the one used above for H_x , we can show that $E_x = 0$ during the entire observation. This leaves a set of only two equations involving E_y and H_z . We designate the mode determined by this set as an x -directed, y -polarized TEM wave in one dimension:

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon} \left[-\frac{\partial H_z}{\partial x} - (J_{\text{source}_y} + \sigma E_y) \right] \quad (3.18a)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left[-\frac{\partial E_y}{\partial x} - (M_{\text{source}_z} + \sigma^* H_z) \right] \quad (3.18b)$$

3.5 EQUIVALENCE TO THE WAVE EQUATION IN ONE DIMENSION

Consider the one-dimensional TEM mode given by (3.16). From this mode, we now derive a homogeneous, lossless, one-dimensional scalar wave equation for H_y , assuming that $M_{\text{source}_y} = J_{\text{source}_z} = 0$ and $\sigma^* = \sigma = 0$. First, take the partial time derivative of (3.16a):

$$\frac{\partial}{\partial t} \left(\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \cdot \frac{\partial E_z}{\partial x} \right) \rightarrow \frac{\partial^2 H_y}{\partial t^2} = \frac{1}{\mu} \cdot \frac{\partial^2 E_z}{\partial t \partial x} \quad (3.19a)$$

Now, take the partial space derivative of (3.16b):

$$\frac{\partial}{\partial x} \left(\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \cdot \frac{\partial H_y}{\partial x} \right) \rightarrow \frac{\partial^2 E_z}{\partial x \partial t} = \frac{1}{\epsilon} \cdot \frac{\partial^2 H_y}{\partial x^2} \quad (3.19b)$$

Since the order of partial differentiation is irrelevant because of the linearity of the system, we substitute the x - t derivative of E_z in (3.19b) into the t - x derivative of E_z in (3.19a) to yield

$$\frac{\partial^2 H_y}{\partial t^2} = \frac{1}{\mu} \cdot \frac{1}{\epsilon} \cdot \frac{\partial^2 H_y}{\partial x^2} = c^2 \frac{\partial^2 H_y}{\partial x^2} \quad (3.19c)$$

where $c = 1/\sqrt{\mu\epsilon}$. Equation (3.19c) is a one-dimensional scalar wave equation for H_y , which occupies the role of u in (2.1). The proportionality factor c , which we earlier showed equals the phase and group velocities of the propagating waves that are solutions to the wave equation, is related to the permeability and permittivity of the medium. For free space wherein $\mu = \mu_0$ and $\epsilon = \epsilon_0$, $c \approx 3 \times 10^8$ meters/second, the speed of light in vacuum.

To obtain the wave equation for E_z , we take the partial time derivative of (3.16b), again assuming that all sources and loss terms equal zero:

$$\frac{\partial}{\partial t} \left(\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \cdot \frac{\partial H_y}{\partial x} \right) \rightarrow \frac{\partial^2 E_z}{\partial t^2} = \frac{1}{\epsilon} \cdot \frac{\partial^2 H_y}{\partial t \partial x} \quad (3.20a)$$

Now take the partial space derivative of (3.16a):

$$\frac{\partial}{\partial x} \left(\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \cdot \frac{\partial E_z}{\partial x} \right) \rightarrow \frac{\partial^2 H_y}{\partial x \partial t} = \frac{1}{\mu} \cdot \frac{\partial^2 E_z}{\partial x^2} \quad (3.20b)$$

Upon substituting the x - t derivative of H_y in (3.20b) into the t - x derivative of H_y in (3.20a), we obtain

$$\frac{\partial^2 E_z}{\partial t^2} = \frac{1}{\epsilon} \cdot \frac{1}{\mu} \cdot \frac{\partial^2 E_z}{\partial x^2} = c^2 \frac{\partial^2 E_z}{\partial x^2} \quad (3.20c)$$

where again $c = 1/\sqrt{\mu\epsilon}$. Equation (3.20c) is a one-dimensional scalar wave equation for E_z , which occupies the role of u in (2.1). The identical proportionality factor c that appeared in (3.19c) appears here as well. Therefore, it is clear that the one-dimensional TEM mode of (3.16) provides for propagating waves of E - and H -components that travel at c . It is left to the student to show that exactly the same results are obtained for the one-dimensional TEM mode of (3.18).

3.6 THE YEE ALGORITHM

3.6.1 Basic Ideas

In 1966, Kane Yee originated a set of finite-difference equations for the time-dependent Maxwell's curl equations system of (3.9) and (3.10) for the lossless materials case $\sigma^* = 0$ and $\sigma = 0$ [1]. Yee's algorithm, introduced in this section, persists in having great usefulness, since its fundamental basis is so robust. Namely:

1. The Yee algorithm solves for *both* electric and magnetic fields in time and space using the coupled Maxwell's curl equations rather than solving for the electric field alone (or the magnetic field alone) with a wave equation.
 - This is analogous to the combined-field integral equation formulation of MM, wherein both \vec{E} and \vec{H} boundary conditions are enforced on the surface of a material structure.
 - Using both \vec{E} and \vec{H} information, the solution is more robust than using either alone (i.e., it is accurate for a wider class of structures). Both electric and magnetic material properties can be modeled in a straightforward manner. This is especially important when modeling radar cross section mitigation.
 - Features unique to each field such as tangential \vec{H} singularities near edges and corners, azimuthal (looping) \vec{H} singularities near thin wires, and radial \vec{E} singularities near points, edges, and thin wires can be individually modeled if both electric and magnetic fields are available.
2. As illustrated in Fig. 3.1, the Yee algorithm centers its \vec{E} and \vec{H} components in three-dimensional space so that every \vec{E} component is surrounded by four circulating \vec{H} components, and every \vec{H} component is surrounded by four circulating \vec{E} components.

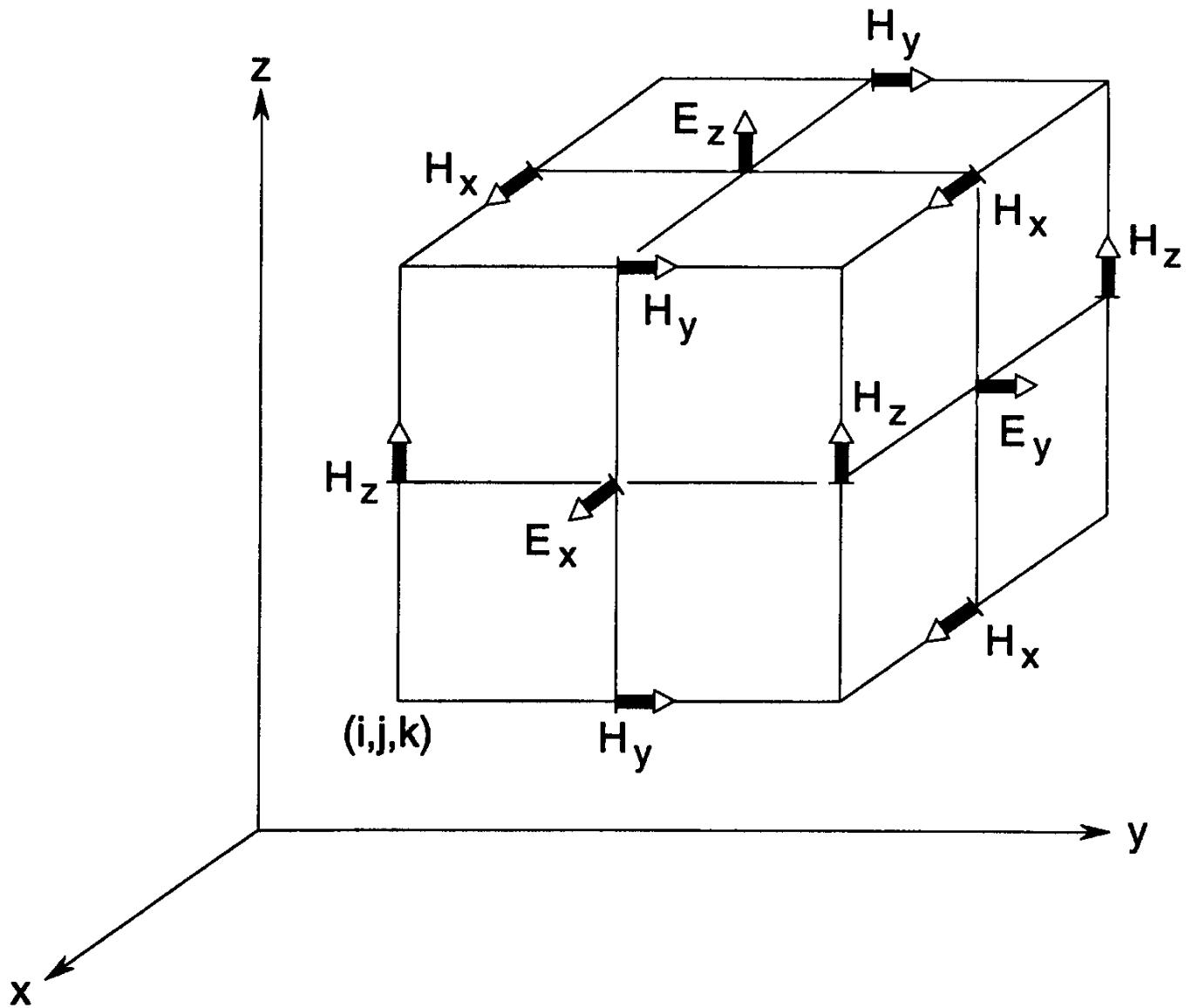


Fig. 3.1 Position of the electric and magnetic field vector components about a cubic unit cell of the Yee space lattice. After: K. S. Yee, *IEEE Trans. Antennas and Propagation*, Vol. 14, 1966, pp. 302–307, © 1966 IEEE.

This provides a beautifully simple picture of three-dimensional space being filled by an interlinked array of Faraday's Law and Ampere's Law contours. For example, it is possible to identify Yee \vec{E} components associated with displacement current flux linking \vec{H} loops, as well as \vec{H} components associated with magnetic flux linking \vec{E} loops. In effect, the Yee algorithm simultaneously simulates the pointwise differential form *and* the macroscopic integral form of Maxwell's equations. The latter is extremely useful in specifying field boundary conditions and singularities.

In addition, we have the following attributes of the Yee space lattice:

- The finite-difference expressions for the space derivatives used in the curl operators are central-difference in nature and second-order accurate.
- Continuity of tangential \vec{E} and \vec{H} is naturally maintained across an interface of dissimilar materials if the interface is parallel to one of the lattice coordinate axes. For this case, there is no need to specially enforce field boundary conditions at the interface. At the beginning of the problem, we simply specify the material permittivity and permeability at each field component location. This yields a stepped or “staircase” approximation of the surface and internal geometry of the structure, with a space resolution set by the size of the lattice unit cell.
- The location of the \vec{E} and \vec{H} components in the Yee space lattice and the central-difference operations on these components implicitly enforce the two Gauss' Law relations (see Section 3.6.9). Thus, the Yee mesh is divergence-free with respect to its E and H fields in the absence of free electric and magnetic charge.

3. As illustrated in Fig. 3.2, the Yee algorithm also centers its \vec{E} and \vec{H} components in time in what is termed a leapfrog arrangement. All of the \vec{E} computations in the modeled space are completed and stored in memory for a particular time point using previously stored \vec{H} data. Then all of the \vec{H} computations in the space are completed and stored in memory using the \vec{E} data just computed. The cycle begins again with the recomputation of the \vec{E} components based on the newly obtained \vec{H} . This process continues until time-stepping is concluded.

- Leapfrog time-stepping is fully explicit, thereby avoiding problems involved with simultaneous equations and matrix inversion.
- The finite-difference expressions for the time derivatives are central-difference in nature and second-order accurate.
- The time-stepping algorithm is nondissipative. That is, numerical wave modes propagating in the mesh do not spuriously decay due to a nonphysical artifact of the time-stepping algorithm.

3.6.2 Finite Differences and Notation

Yee introduced the notation used in Section 2.1 (there, in one spatial dimension) for space points and functions of space and time. For convenience, this notation is repeated here and generalized to three spatial dimensions. We denote a space point in a uniform, rectangular lattice as

$$(i, j, k) = (i\Delta x, j\Delta y, k\Delta z) \quad (3.21)$$

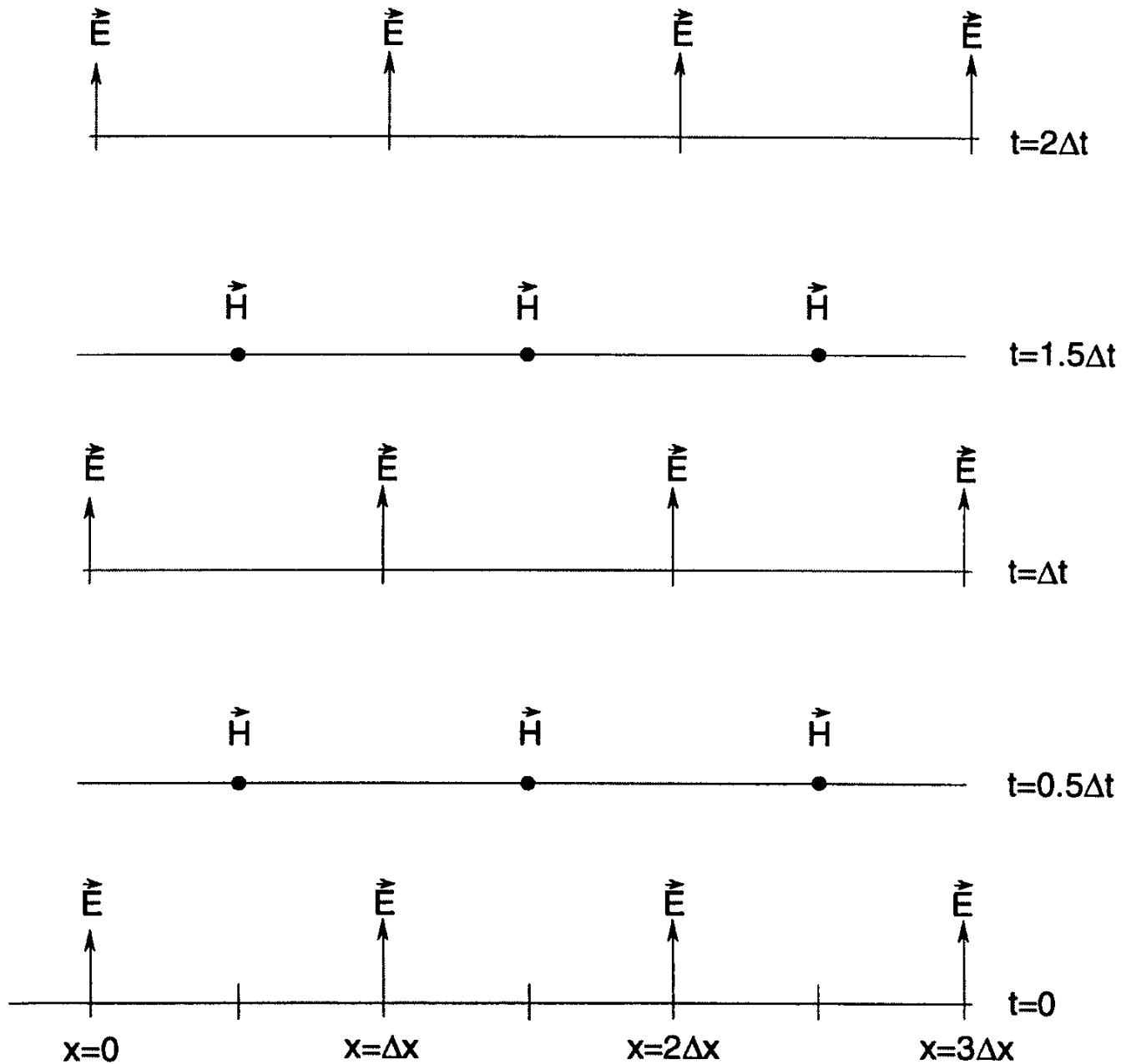


Fig. 3.2 Space-time chart of the Yee algorithm for a one-dimensional wave propagation example showing the use of central differences for the space derivatives and leapfrog for the time derivatives. Initial conditions for both electric and magnetic fields are zero everywhere in the grid.

Here, Δx , Δy , and Δz are, respectively, the lattice space increments in the x , y , and z coordinate directions, and i , j , and k are integers. Further, we denote any function u of space and time evaluated at a discrete point in the grid and at a discrete point in time as

$$u(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = u_{i,j,k}^n \quad (3.22)$$

where Δt is the time increment, assumed uniform over the observation interval, and n is an integer.

Yee used centered finite-difference (central-difference) expressions for the space and time derivatives that are both simply programmed and second-order accurate in the space and time increments. Consider his expression for the first partial space derivative of u in the x -direction, evaluated at the fixed time $t_n = n\Delta t$:

$$\frac{\partial u}{\partial x}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i+1/2,j,k}^n - u_{i-1/2,j,k}^n}{\Delta x} + O[(\Delta x)^2] \quad (3.23)$$

We note the $\pm 1/2$ increment in the i subscript (x -coordinate) of u , denoting a space finite-difference over $\pm 1/2 \Delta x$. Recalling the tutorial material on finite differences in Section 2.1, it is clear that Yee derived (3.23) by subtracting (2.10b) from (2.10a) (but with $\Delta x/2$ substituted for Δx in these expressions), and then solving for $\partial u / \partial x$. Yee's goal was second-order accurate central differencing, but it is apparent that he desired to take data for his central differences to the right and left of his observation point by only $\Delta x/2$, rather than a full Δx .

Yee chose this notation because he wished to interleave his \vec{E} and \vec{H} components in the space lattice at intervals of $\Delta x/2$. For example, the difference of two adjacent \vec{E} components, separated by Δx and located $\pm 1/2 \Delta x$ on either side of an \vec{H} component, would be used to provide a numerical approximation for $\partial E / \partial x$ to permit stepping the \vec{H} component in time. For completeness, it should be added that a numerical approximation analogous to (3.23) for $\partial u / \partial y$ or $\partial u / \partial z$ can be written simply by incrementing the j or k subscript of u by $\pm 1/2 \Delta y$ or $\pm 1/2 \Delta z$, respectively.

Yee's expression for the first time partial derivative of u , evaluated at the fixed space point (i, j, k) , follows by analogy:

$$\frac{\partial u}{\partial t}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{u_{i,j,k}^{n+1/2} - u_{i,j,k}^{n-1/2}}{\Delta t} + O[(\Delta t)^2] \quad (3.24)$$

Now the $\pm 1/2$ increment is in the n superscript (time coordinate) of u , denoting a time finite-difference over $\pm 1/2 \Delta t$. Yee chose this notation because he wished to interleave his \vec{E} and \vec{H} components in time at intervals of $1/2 \Delta t$ for purposes of implementing a leapfrog algorithm.

3.6.3 Finite-Difference Expressions for Maxwell's Equations in Three Dimensions

We now apply the above ideas and notation to achieve a numerical approximation of the Maxwell's curl equations in three dimensions given by the system of equations of (3.9) and (3.10). We begin by considering the E_x field-component equation (3.10a), repeated here for convenience:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \left[\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - (J_{\text{source}_x} + \sigma E_x) \right] \quad (3.10a)$$

Referring to Fig. 3.1, consider a typical substitution of central differences for the time and space derivatives in (3.10a), for example at $E_x(i, j+1/2, k+1/2, n)$. Here, we have initially

$$\begin{aligned} \frac{E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j+1/2,k+1/2}^{n-1/2}}{\Delta t} &= \\ \frac{1}{\epsilon_{i,j+1/2,k+1/2}} \cdot &\left(\begin{array}{l} \frac{H_z|_{i,j+1,k+1/2}^n - H_z|_{i,j,k+1/2}^n}{\Delta y} - \frac{H_y|_{i,j+1/2,k+1}^n - H_y|_{i,j+1/2,k}^n}{\Delta z} \\ - J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n - \sigma_{i,j+1/2,k+1/2} E_x|_{i,j+1/2,k+1/2}^n \end{array} \right) \end{aligned} \quad (3.25)$$

Note that all field quantities on the right-hand side are evaluated at time-step n , including the electric field E_x appearing due to the material conductivity σ . Since E_x values at time-step n are not assumed to be stored in the computer's memory (only the previous values of E_x at time-step $n-1/2$ are assumed to be in memory), we need some way to estimate such terms. A very good way is as follows, using what we call a *semi-implicit approximation*:

$$E_x|_{i,j+1/2,k+1/2}^n = \frac{E_x|_{i,j+1/2,k+1/2}^{n+1/2} + E_x|_{i,j+1/2,k+1/2}^{n-1/2}}{2} \quad (3.26)$$

Here E_x values at time-step n are assumed to be simply the arithmetic average of the stored values of E_x at time-step $n-1/2$ and the yet-to-be computed new values of E_x .

at time-step $n + 1/2$. Substituting (3.26) into (3.25) after multiplying both sides by Δt , we obtain

$$\frac{E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j+1/2,k+1/2}^{n-1/2}}{\Delta t} = \frac{H_z|_{i,j+1/2,k+1/2}^n - H_z|_{i,j,k+1/2}^n}{\Delta y} - \frac{H_y|_{i,j+1/2,k+1}^n - H_y|_{i,j+1/2,k}^n}{\Delta z} - \frac{J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n - \sigma_{i,j+1/2,k+1/2} \cdot \left(\frac{E_x|_{i,j+1/2,k+1/2}^{n+1/2} + E_x|_{i,j+1/2,k+1/2}^{n-1/2}}{2} \right)}{\epsilon_{i,j+1/2,k+1/2}} \quad (3.27)$$

We note that the terms $E_x|_{i,j+1/2,k+1/2}^{n+1/2}$ and $E_x|_{i,j+1/2,k+1/2}^{n-1/2}$ appear on both sides of (3.27). Collecting all terms of these two types and isolating $E_x|_{i,j+1/2,k+1/2}^{n+1/2}$ on the left-hand side yields

$$\begin{aligned} \left(1 + \frac{\sigma_{i,j+1/2,k+1/2} \Delta t}{2 \epsilon_{i,j+1/2,k+1/2}} \right) E_x|_{i,j+1/2,k+1/2}^{n+1/2} &= \left(1 - \frac{\sigma_{i,j+1/2,k+1/2} \Delta t}{2 \epsilon_{i,j+1/2,k+1/2}} \right) E_x|_{i,j+1/2,k+1/2}^{n-1/2} \\ &+ \frac{\Delta t}{\epsilon_{i,j+1/2,k+1/2}} \cdot \left(\frac{H_z|_{i,j+1,k+1/2}^n - H_z|_{i,j,k+1/2}^n}{\Delta y} - \frac{H_y|_{i,j+1/2,k+1}^n - H_y|_{i,j+1/2,k}^n}{\Delta z} \right. \\ &\quad \left. - J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n \right) \end{aligned} \quad (3.28)$$

Dividing both sides by $(1 + \sigma_{i,j+1/2,k+1/2} \Delta t / 2 \epsilon_{i,j+1/2,k+1/2})$ yields the desired explicit time-stepping relation for $E_x|_{i,j+1/2,k+1/2}^{n+1/2}$:

$$\begin{aligned}
E_x|_{i, j+1/2, k+1/2}^{n+1/2} &= \left(\frac{1 - \frac{\sigma_{i, j+1/2, k+1/2} \Delta t}{2 \epsilon_{i, j+1/2, k+1/2}}}{1 + \frac{\sigma_{i, j+1/2, k+1/2} \Delta t}{2 \epsilon_{i, j+1/2, k+1/2}}} \right) E_x|_{i, j+1/2, k+1/2}^{n-1/2} \\
&+ \left(\frac{\frac{\Delta t}{\epsilon_{i, j+1/2, k+1/2}}}{1 + \frac{\sigma_{i, j+1/2, k+1/2} \Delta t}{2 \epsilon_{i, j+1/2, k+1/2}}} \right) \cdot \left(\begin{array}{l} \frac{H_z|_{i, j+1, k+1/2}^n - H_z|_{i, j, k+1/2}^n}{\Delta y} \\ - \frac{H_y|_{i, j+1/2, k+1}^n - H_y|_{i, j+1/2, k}^n}{\Delta z} \\ - J_{\text{source}_x}|_{i, j+1/2, k+1/2}^n \end{array} \right)
\end{aligned} \tag{3.29a}$$

The semi-implicit assumption of (3.26) has been found to yield numerically stable and accurate results for values of σ from zero to infinity. As we have seen above, this assumption fortunately allows us to avoid simultaneous equations for $E_x|^{n+1/2}$. The term of this type introduced on the right-hand side of (3.27) can be grouped with a like term on the left-hand side and then solved explicitly.

Similarly, we can derive finite-difference expressions based on Yee's algorithm for the E_y and E_z field components given by Maxwell's equations (3.10b) and (3.10c). Referring again to Fig. 3.1, we have for example the following time-stepping expressions for the E components normal to the remaining visible faces of the unit cell:

$$\begin{aligned}
E_y|_{i-1/2, j+1, k+1/2}^{n+1/2} &= \left(\frac{1 - \frac{\sigma_{i-1/2, j+1, k+1/2} \Delta t}{2 \epsilon_{i-1/2, j+1, k+1/2}}}{1 + \frac{\sigma_{i-1/2, j+1, k+1/2} \Delta t}{2 \epsilon_{i-1/2, j+1, k+1/2}}} \right) E_y|_{i-1/2, j+1, k+1/2}^{n-1/2} \\
&+ \left(\frac{\frac{\Delta t}{\epsilon_{i-1/2, j+1, k+1/2}}}{1 + \frac{\sigma_{i-1/2, j+1, k+1/2} \Delta t}{2 \epsilon_{i-1/2, j+1, k+1/2}}} \right) \cdot \left(\begin{array}{l} \frac{H_x|_{i-1/2, j+1, k+1}^n - H_x|_{i-1/2, j+1, k}^n}{\Delta z} \\ - \frac{H_z|_{i, j+1, k+1/2}^n - H_z|_{i-1, j+1, k+1/2}^n}{\Delta x} \\ - J_{\text{source}_y}|_{i-1/2, j+1, k+1/2}^n \end{array} \right)
\end{aligned} \tag{3.29b}$$

$$E_z|_{i-1/2, j+1/2, k+1}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i-1/2, j+1/2, k+1} \Delta t}{2 \epsilon_{i-1/2, j+1/2, k+1}}}{1 + \frac{\sigma_{i-1/2, j+1/2, k+1} \Delta t}{2 \epsilon_{i-1/2, j+1/2, k+1}}} \right) E_z|_{i-1/2, j+1/2, k+1}^{n-1/2} \\ + \left(\frac{\frac{\Delta t}{\epsilon_{i-1/2, j+1/2, k+1}}}{1 + \frac{\sigma_{i-1/2, j+1/2, k+1} \Delta t}{2 \epsilon_{i-1/2, j+1/2, k+1}}} \right) \cdot \begin{cases} \frac{H_y|_{i, j+1/2, k+1}^n - H_y|_{i-1, j+1/2, k+1}^n}{\Delta x} \\ - \frac{H_x|_{i-1/2, j+1, k+1}^n - H_x|_{i-1/2, j, k+1}^n}{\Delta y} \\ - J_{\text{source}_z}|_{i-1/2, j+1/2, k+1}^n \end{cases} \quad (3.29c)$$

By analogy we can derive finite-difference equations for (3.9a)–(3.9c) to time-step H_x , H_y , and H_z . Here $\sigma^* H$ represents a magnetic loss term on the right-hand side of each equation, which is estimated using a semi-implicit procedure analogous to (3.26). This results in three equations having a form similar to that of the E equations above. Referring again to Fig. 3.1, we have for example the following time-stepping expression for the H_x component located at the upper right corner of the unit cell:

$$H_x|_{i-1/2, j+1, k+1}^{n+1} = \left(\frac{1 - \frac{\sigma_{i-1/2, j+1, k+1}^* \Delta t}{2 \mu_{i-1/2, j+1, k+1}}}{1 + \frac{\sigma_{i-1/2, j+1, k+1}^* \Delta t}{2 \mu_{i-1/2, j+1, k+1}}} \right) H_x|_{i-1/2, j+1, k+1}^n \\ + \left(\frac{\frac{\Delta t}{\mu_{i-1/2, j+1, k+1}}}{1 + \frac{\sigma_{i-1/2, j+1, k+1}^* \Delta t}{2 \mu_{i-1/2, j+1, k+1}}} \right) \cdot \begin{cases} \frac{E_y|_{i-1/2, j+1, k+3/2}^{n+1/2} - E_y|_{i-1/2, j+1, k+1/2}^{n+1/2}}{\Delta z} \\ - \frac{E_z|_{i-1/2, j+3/2, k+1}^{n+1/2} - E_z|_{i-1/2, j+1/2, k+1}^{n+1/2}}{\Delta y} \\ - M_{\text{source}_x}|_{i-1/2, j+1, k+1}^{n+1/2} \end{cases} \quad (3.30a)$$

Similarly, we have the following time-stepping expression for the H_y component located at the upper front corner of the unit cell:

$$H_y|_{i, j+1/2, k+1}^{n+1} = \left(\frac{1 - \frac{\sigma^*_{i, j+1/2, k+1} \Delta t}{2 \mu_{i, j+1/2, k+1}}}{1 + \frac{\sigma^*_{i, j+1/2, k+1} \Delta t}{2 \mu_{i, j+1/2, k+1}}} \right) H_y|_{i, j+1/2, k+1}^n$$

$$+ \left(\frac{\frac{\Delta t}{\mu_{i, j+1/2, k+1}}}{1 + \frac{\sigma^*_{i, j+1/2, k+1} \Delta t}{2 \mu_{i, j+1/2, k+1}}} \right) \cdot \begin{cases} \frac{E_z|_{i+1/2, j+1/2, k+1}^{n+1/2} - E_z|_{i-1/2, j+1/2, k+1}^{n+1/2}}{\Delta x} \\ - \frac{E_x|_{i, j+1/2, k+3/2}^{n+1/2} - E_x|_{i, j+1/2, k+1/2}^{n+1/2}}{\Delta z} \\ - M_{\text{source}_y}|_{i, j+1/2, k+1}^{n+1/2} \end{cases} \quad (3.30b)$$

Finally, we have the following time-stepping expression for the H_z component located at the right front corner of the unit cell:

$$H_z|_{i, j+1, k+1/2}^{n+1} = \left(\frac{1 - \frac{\sigma^*_{i, j+1, k+1/2} \Delta t}{2 \mu_{i, j+1, k+1/2}}}{1 + \frac{\sigma^*_{i, j+1, k+1/2} \Delta t}{2 \mu_{i, j+1, k+1/2}}} \right) H_z|_{i, j+1, k+1/2}^n$$

$$+ \left(\frac{\frac{\Delta t}{\mu_{i, j+1, k+1/2}}}{1 + \frac{\sigma^*_{i, j+1, k+1/2} \Delta t}{2 \mu_{i, j+1, k+1/2}}} \right) \cdot \begin{cases} \frac{E_x|_{i, j+3/2, k+1/2}^{n+1/2} - E_x|_{i, j+1/2, k+1/2}^{n+1/2}}{\Delta y} \\ - \frac{E_y|_{i+1/2, j+1, k+1/2}^{n+1/2} - E_y|_{i-1/2, j+1, k+1/2}^{n+1/2}}{\Delta x} \\ - M_{\text{source}_z}|_{i, j+1, k+1/2}^{n+1/2} \end{cases} \quad (3.30c)$$

With the systems of finite-difference expressions of (3.29) and (3.30), the new value of an electromagnetic field vector component at any lattice point depends only on its previous value, the previous values of the components of the other field vector at adjacent points, and the known electric and magnetic current sources. Therefore, at any given time step, the computation of a field vector can proceed either one point at a time, or, if p parallel processors are employed concurrently, p points at a time.

3.6.4 Space Region With a Continuous Variation of Material Properties

To implement the finite-difference systems of (3.29) and (3.30) for a region having a continuous variation of material properties with spatial position, it is desirable to define and store the following constant updating coefficients for each field vector component before the time-stepping begins:

Updating Coefficients at the General E-Field Component Location (i, j, k):

$$C_a|_{i,j,k} = \left(1 - \frac{\sigma_{i,j,k} \Delta t}{2\epsilon_{i,j,k}} \right) / \left(1 + \frac{\sigma_{i,j,k} \Delta t}{2\epsilon_{i,j,k}} \right) \quad (3.31a)$$

$$C_{b_1}|_{i,j,k} = \left(\frac{\Delta t}{\epsilon_{i,j,k} \Delta_1} \right) / \left(1 + \frac{\sigma_{i,j,k} \Delta t}{2\epsilon_{i,j,k}} \right) \quad (3.31b)$$

$$C_{b_2}|_{i,j,k} = \left(\frac{\Delta t}{\epsilon_{i,j,k} \Delta_2} \right) / \left(1 + \frac{\sigma_{i,j,k} \Delta t}{2\epsilon_{i,j,k}} \right) \quad (3.31c)$$

Updating Coefficients at the General H-Field Component Location (i, j, k):

$$D_a|_{i,j,k} = \left(1 - \frac{\sigma^*_{i,j,k} \Delta t}{2\mu_{i,j,k}} \right) / \left(1 + \frac{\sigma^*_{i,j,k} \Delta t}{2\mu_{i,j,k}} \right) \quad (3.32a)$$

$$D_{b_1}|_{i,j,k} = \left(\frac{\Delta t}{\mu_{i,j,k} \Delta_1} \right) / \left(1 + \frac{\sigma^*_{i,j,k} \Delta t}{2\mu_{i,j,k}} \right) \quad (3.32b)$$

$$D_{b_2}|_{i,j,k} = \left(\frac{\Delta t}{\mu_{i,j,k} \Delta_2} \right) / \left(1 + \frac{\sigma^*_{i,j,k} \Delta t}{2\mu_{i,j,k}} \right) \quad (3.32c)$$

In (3.31) and (3.32), Δ_1 and Δ_2 denote the two possible lattice space increments used for the finite differences in each field-component calculation. For a cubic lattice, $\Delta x = \Delta y = \Delta z = \Delta$ and thus $\Delta_1 = \Delta_2 = \Delta$. For this case, $C_{b_1} = C_{b_2}$ and $D_{b_1} = D_{b_2}$, reducing the storage requirement to two updating coefficients per field vector component. For this case, the approximate total computer storage needed is $18N$, where N is the number of space cells in the FDTD lattice. The finite-difference expressions of (3.29) can now be rewritten more simply as:

$$E_x|_{i,j+1/2,k+1/2}^{n+1/2} = C_{a,E_x}|_{i,j+1/2,k+1/2} E_x|_{i,j+1/2,k+1/2}^{n-1/2} \quad (3.33a)$$

$$+ C_{b,E_x}|_{i,j+1/2,k+1/2} \cdot \begin{pmatrix} H_z|_{i,j+1,k+1/2}^n - H_z|_{i,j,k+1/2}^n + \\ H_y|_{i,j+1/2,k}^n - H_y|_{i,j+1/2,k+1}^n - J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n \Delta \end{pmatrix}$$

$$E_y|_{i-1/2,j+1,k+1/2}^{n+1/2} = C_{a,E_y}|_{i-1/2,j+1,k+1/2} E_y|_{i-1/2,j+1,k+1/2}^{n-1/2} \quad (3.33b)$$

$$+ C_{b,E_y}|_{i-1/2,j+1,k+1/2} \cdot \begin{pmatrix} H_x|_{i-1/2,j+1,k+1}^n - H_x|_{i-1/2,j+1,k}^n + \\ H_z|_{i-1,j+1,k+1/2}^n - H_z|_{i,j+1,k+1/2}^n - J_{\text{source}_y}|_{i-1/2,j+1,k+1/2}^n \Delta \end{pmatrix}$$

$$E_z|_{i-1/2,j+1/2,k+1}^{n+1/2} = C_{a,E_z}|_{i-1/2,j+1/2,k+1} E_z|_{i-1/2,j+1/2,k+1}^{n-1/2} \quad (3.33c)$$

$$+ C_{b,E_z}|_{i-1/2,j+1/2,k+1} \cdot \begin{pmatrix} H_y|_{i,j+1/2,k+1}^n - H_y|_{i-1,j+1/2,k+1}^n + \\ H_x|_{i-1/2,j,k+1}^n - H_x|_{i-1/2,j+1,k+1}^n - J_{\text{source}_z}|_{i-1/2,j+1/2,k+1}^n \Delta \end{pmatrix}$$

The finite-difference expressions of (3.30) can now be rewritten more simply as:

$$H_x|_{i-1/2,j+1,k+1}^{n+1} = D_{a,H_x}|_{i-1/2,j+1,k+1} H_x|_{i-1/2,j+1,k+1}^n \quad (3.34a)$$

$$+ D_{b,H_x}|_{i-1/2,j+1,k+1} \cdot \begin{pmatrix} E_y|_{i-1/2,j+1,k+3/2}^{n+1/2} - E_y|_{i-1/2,j+1,k+1/2}^{n+1/2} + \\ E_z|_{i-1/2,j+1/2,k+1}^{n+1/2} - E_z|_{i-1/2,j+3/2,k+1}^{n+1/2} - M_{\text{source}_x}|_{i-1/2,j+1,k+1}^{n+1/2} \Delta \end{pmatrix}$$

$$H_y|_{i,j+1/2,k+1}^{n+1} = D_{a,H_y}|_{i,j+1/2,k+1} H_y|_{i,j+1/2,k+1}^n \quad (3.34b)$$

$$+ D_{b,H_y}|_{i,j+1/2,k+1} \cdot \begin{pmatrix} E_z|_{i+1/2,j+1/2,k+1}^{n+1/2} - E_z|_{i-1/2,j+1/2,k+1}^{n+1/2} + \\ E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j+1/2,k+3/2}^{n+1/2} - M_{\text{source}_y}|_{i,j+1/2,k+1}^{n+1/2} \Delta \end{pmatrix}$$

$$H_z|_{i,j+1,k+1/2}^{n+1} = D_{a,H_z}|_{i,j+1,k+1/2} H_z|_{i,j+1,k+1/2}^n + D_{b,H_z}|_{i,j+1,k+1/2} \cdot \left(\begin{array}{l} E_x|_{i,j+3/2,k+1/2}^{n+1/2} - E_x|_{i,j+1/2,k+1/2}^{n+1/2} + \\ E_y|_{i-1/2,j+1,k+1/2}^{n+1/2} - E_y|_{i+1/2,j+1,k+1/2}^{n+1/2} - M_{\text{source}_z}|_{i,j+1,k+1/2}^{n+1/2} \Delta \end{array} \right) \quad (3.34c)$$

3.6.5 Space Region With a Finite Number of Distinct Media

For a space region with a finite number of media having distinct electrical properties, the computer storage requirement can be further reduced. This can be done by defining an integer array, $\text{MEDIA}(i, j, k)$, for each set of field vector components. This array stores an integer “pointer” at each location of such a field component in the space lattice, enabling the proper algorithm coefficients to be extracted. For this case, the finite-difference expressions of (3.33) can be rewritten more simply as:

$$m = \text{MEDIA}_{E_x}|_{i,j+1/2,k+1/2} \quad (3.35a)$$

$$E_x|_{i,j+1/2,k+1/2}^{n+1/2} = C_a(m) E_x|_{i,j+1/2,k+1/2}^{n-1/2} + C_b(m) \cdot \left(H_z|_{i,j+1,k+1/2}^n - H_z|_{i,j,k+1/2}^n + H_y|_{i,j+1/2,k}^n - H_y|_{i,j+1/2,k+1}^n - J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n \Delta \right)$$

$$m = \text{MEDIA}_{E_y}|_{i-1/2,j+1,k+1/2} \quad (3.35b)$$

$$E_y|_{i-1/2,j+1,k+1/2}^{n+1/2} = C_a(m) E_y|_{i-1/2,j+1,k+1/2}^{n-1/2} + C_b(m) \cdot \left(H_x|_{i-1/2,j+1,k+1}^n - H_x|_{i-1/2,j+1,k}^n + H_z|_{i-1,j+1,k+1/2}^n - H_z|_{i,j+1,k+1/2}^n - J_{\text{source}_y}|_{i-1/2,j+1,k+1/2}^n \Delta \right)$$

$$m = \text{MEDIA}_{E_z}|_{i-1/2,j+1/2,k+1} \quad (3.35c)$$

$$E_z|_{i-1/2,j+1/2,k+1}^{n+1/2} = C_a(m) E_z|_{i-1/2,j+1/2,k+1}^{n-1/2} + C_b(m) \cdot \left(H_y|_{i,j+1/2,k+1}^n - H_y|_{i-1,j+1/2,k+1}^n + H_x|_{i-1/2,j,k+1}^n - H_x|_{i-1/2,j+1,k+1}^n - J_{\text{source}_z}|_{i-1/2,j+1/2,k+1}^n \Delta \right)$$

The finite-difference expressions of (3.34) can be rewritten more simply as:

$$m = \text{MEDIA}_{H_x}|_{i-1/2, j+1, k+1} \quad (3.36a)$$

$$H_x|_{i-1/2, j+1, k+1}^{n+1} = D_a(m) H_x|_{i-1/2, j+1, k+1}^n + D_b(m) \cdot \left(E_y|_{i-1/2, j+1, k+3/2}^{n+1/2} - E_y|_{i-1/2, j+1, k+1/2}^{n+1/2} + E_z|_{i-1/2, j+1/2, k+1}^{n+1/2} - E_z|_{i-1/2, j+3/2, k+1}^{n+1/2} - M_{\text{source}_x}|_{i-1/2, j+1, k+1}^{n+1/2} \Delta \right)$$

$$m = \text{MEDIA}_{H_y}|_{i, j+1/2, k+1} \quad (3.36b)$$

$$H_y|_{i, j+1/2, k+1}^{n+1} = D_a(m) H_y|_{i, j+1/2, k+1}^n + D_b(m) \cdot \left(E_z|_{i+1/2, j+1/2, k+1}^{n+1/2} - E_z|_{i-1/2, j+1/2, k+1}^{n+1/2} + E_x|_{i, j+1/2, k+1/2}^{n+1/2} - E_x|_{i, j+1/2, k+3/2}^{n+1/2} - M_{\text{source}_y}|_{i, j+1/2, k+1}^{n+1/2} \Delta \right)$$

$$m = \text{MEDIA}_{H_z}|_{i, j+1, k+1/2} \quad (3.36c)$$

$$H_z|_{i, j+1, k+1/2}^{n+1} = D_a(m) H_z|_{i, j+1, k+1/2}^n + D_b(m) \cdot \left(E_x|_{i, j+3/2, k+1/2}^{n+1/2} - E_x|_{i, j+1/2, k+1/2}^{n+1/2} + E_y|_{i-1/2, j+1, k+1/2}^{n+1/2} - E_y|_{i+1/2, j+1, k+1/2}^{n+1/2} - M_{\text{source}_z}|_{i, j+1, k+1/2}^{n+1/2} \Delta \right)$$

With the finite-difference systems of (3.35) and (3.36), we note that the coefficient arrays $C_a(m)$, $C_b(m)$, $D_a(m)$, and $D_b(m)$ each contain only M elements, where M is the number of distinct material media in the FDTD space lattice. Thus, if separate $\text{MEDIA}(i, j, k)$ integer pointer arrays are provided for each field vector component, the approximate total computer storage needed is reduced to $12N$, where N is the number of space cells in the FDTD lattice. This reduction in computer storage comes at some cost, however, since additional computer instructions must be executed at each field-vector location to obtain the pointer integer m from the associated MEDIA array and then extract the $C(m)$ or $D(m)$ updating coefficients. On a vectorizing computer like the Cray, these additional instructions can inhibit the ultrafast flow of vectorized computations possible when no pointer arrays are used, causing a substantial reduction in throughput. However, the latter can be mitigated through careful programming.

Taking advantage of the integer nature of the MEDIA arrays, further reduction in computer storage can be achieved. For example, FDTD codes have been written that use word packing to combine the floating-point field value at each space lattice point with its associated MEDIA integer. This halves the computer storage to $6N$. However, with the additional computer instructions required for word packing and word unpacking, there is a tradeoff between storage and running time for such codes. Word packing would be pursued primarily if going “out of core” leads to an intolerable expansion of program running time due to massive input/output (I/O) to the disk array.

A more efficient means of packing the MEDIA integers is to construct a separate bit-packed array. For example, a 64-bit word can be divided into sixteen 4-bit pointers. Such a composite pointer could specify up to $2^4 = 16$ distinct media at each of 16 field component locations in the grid. This provides the means to reduce the overall computer storage for the MEDIA arrays by a factor of $15/16$ (94%). Efficient vectorized routines for word packing and word unpacking of this type are available on the Cray, and appeared in at least one widely used FDTD code (Lawrence Livermore National Laboratory’s *TSAR*).

3.6.6 Space Region With Nonpermeable Media

Many electromagnetic wave interaction problems involve nonpermeable media ($\mu = \mu_0$, $\sigma^* = 0$) and can be implemented on a uniform cubic-cell FDTD space lattice. For such problems, the finite-difference expressions of (3.35) and (3.36) can be further simplified by defining the proportional $\hat{\vec{E}}$ and $\hat{\vec{M}}$ vectors:

$$\hat{\vec{E}} = (\Delta t / \mu_0 \Delta) \vec{E} ; \quad \hat{\vec{M}} = (\Delta t / \mu_0) \vec{M} \quad (3.37a, b)$$

where $\Delta = \Delta x = \Delta y = \Delta z$ is the cell size of the space lattice. Assuming that \hat{E}_x , \hat{E}_y , and \hat{E}_z are stored in the computer memory, and further defining a scaled E -field updating coefficient $\hat{C}_b(m)$ as

$$\hat{C}_b(m) = (\Delta t / \mu_0 \Delta) C_b(m) \quad (3.38)$$

the finite-difference expressions of (3.35) can be rewritten as:

$$m = \text{MEDIA}_{E_x|_{i,j+1/2,k+1/2}} \quad (3.39a)$$

$$\begin{aligned} \hat{E}_x|_{i,j+1/2,k+1/2}^{n+1/2} &= C_a(m) \hat{E}_x|_{i,j+1/2,k+1/2}^{n-1/2} + \hat{C}_b(m) \cdot \left(H_z|_{i,j+1,k+1/2}^n - \right. \\ &\quad \left. H_z|_{i,j,k+1/2}^n + H_y|_{i,j+1/2,k}^n - H_y|_{i,j+1/2,k+1}^n - J_{\text{source}_x}|_{i,j+1/2,k+1/2}^n \Delta \right) \end{aligned}$$

$$m = \text{MEDIA}_{E_y} \Big|_{i-1/2, j+1, k+1/2} \quad (3.39b)$$

$$\begin{aligned} \hat{E}_y \Big|_{i-1/2, j+1, k+1/2}^{n+1/2} &= C_a(m) \hat{E}_y \Big|_{i-1/2, j+1, k+1/2}^{n-1/2} + \hat{C}_b(m) \cdot \left(H_x \Big|_{i-1/2, j+1, k+1}^n - \right. \\ &\quad \left. H_x \Big|_{i-1/2, j+1, k}^n + H_z \Big|_{i-1, j+1, k+1/2}^n - H_z \Big|_{i, j+1, k+1/2}^n - J_{\text{source}_y} \Big|_{i-1/2, j+1, k+1/2}^n \Delta \right) \end{aligned}$$

$$m = \text{MEDIA}_{E_z} \Big|_{i-1/2, j+1/2, k+1} \quad (3.39c)$$

$$\begin{aligned} \hat{E}_z \Big|_{i-1/2, j+1/2, k+1}^{n+1/2} &= C_a(m) \hat{E}_z \Big|_{i-1/2, j+1/2, k+1}^{n-1/2} + \hat{C}_b(m) \cdot \left(H_y \Big|_{i, j+1/2, k+1}^n - \right. \\ &\quad \left. H_y \Big|_{i-1, j+1/2, k+1}^n + H_x \Big|_{i-1/2, j, k+1}^n - H_x \Big|_{i-1/2, j+1, k+1}^n - J_{\text{source}_z} \Big|_{i-1/2, j+1/2, k+1}^n \Delta \right) \end{aligned}$$

The finite-difference expressions of (3.36) can be rewritten very simply as:

$$H_x \Big|_{i-1/2, j+1, k+1}^{n+1} = H_x \Big|_{i-1/2, j+1, k+1}^n + \hat{E}_y \Big|_{i-1/2, j+1, k+3/2}^{n+1/2} - \quad (3.40a)$$

$$\hat{E}_y \Big|_{i-1/2, j+1, k+1/2}^{n+1/2} + \hat{E}_z \Big|_{i-1/2, j+1/2, k+1}^{n+1/2} - \hat{E}_z \Big|_{i-1/2, j+3/2, k+1}^{n+1/2} - \hat{M}_{\text{source}_x} \Big|_{i-1/2, j+1, k+1}^{n+1/2}$$

$$H_y \Big|_{i, j+1/2, k+1}^{n+1} = H_y \Big|_{i, j+1/2, k+1}^n + \hat{E}_z \Big|_{i+1/2, j+1/2, k+1}^{n+1/2} - \quad (3.40b)$$

$$\hat{E}_z \Big|_{i-1/2, j+1/2, k+1}^{n+1/2} + \hat{E}_x \Big|_{i, j+1/2, k+1/2}^{n+1/2} - \hat{E}_x \Big|_{i, j+1/2, k+3/2}^{n+1/2} - \hat{M}_{\text{source}_y} \Big|_{i, j+1/2, k+1}^{n+1/2}$$

$$H_z \Big|_{i, j+1, k+1/2}^{n+1} = H_z \Big|_{i, j+1, k+1/2}^n + \hat{E}_x \Big|_{i, j+3/2, k+1/2}^{n+1/2} - \quad (3.40c)$$

$$\hat{E}_x \Big|_{i, j+1/2, k+1/2}^{n+1/2} + \hat{E}_y \Big|_{i-1/2, j+1, k+1/2}^{n+1/2} - \hat{E}_y \Big|_{i+1/2, j+1, k+1/2}^{n+1/2} - \hat{M}_{\text{source}_z} \Big|_{i, j+1, k+1/2}^{n+1/2}$$

The simplified finite-difference systems of (3.39) and (3.40) eliminate the three multiplications previously needed to update the H -field components, and require storage of MEDIA arrays only for the E -field components. At the end of the run, the desired values of the unscaled E -fields can be obtained simply by multiplying the stored scaled values by the reciprocal of the scaling factor of (3.37a).

3.6.7 Reduction to the Two-Dimensional TM_z and TE_z Modes

The finite-difference systems of (3.29) and (3.30), (3.33) and (3.34), (3.35) and (3.36), and (3.39) and (3.40) can be reduced to the proper algorithms for the two-dimensional TM_z and TE_z cases summarized in Section 3.3. For convenience and consistency, we again consider the field vector components in the space lattice represented by the unit cell of Fig. 3.1. Assuming now that all partial derivatives of the fields with respect to z are equal to zero, the following conditions hold:

1. The sets of (E_z, H_x, H_y) components located in each lattice cut plane $k, k+1, \dots$ etc. are identical and can be completely represented by any one of these sets, which we designate as the TM_z mode.
2. The sets of (H_z, E_x, E_y) components located in each lattice cut plane $k+1/2, k+3/2, \dots$ etc. are identical and can be completely represented by any one of these sets, which we designate as the TE_z mode.
3. The TM_z and TE_z modes are completely decoupled from each other.

Further assume, for example, a region of space containing a finite number of material media having distinct electrical properties. Then we reduce the finite-difference systems of (3.35) and (3.36) as follows.

TM_z Mode, Corresponding to the System of (3.13)

Write (3.35c), (3.36a), and (3.36b) without specifying the k index, and then set all partial derivatives of the fields with respect to z equal to zero:

$$m = \text{MEDIA}_{E_z} \Big|_{i-1/2, j+1/2}$$

$$E_z|_{i-1/2, j+1/2}^{n+1/2} = C_a(m) E_z|_{i-1/2, j+1/2}^{n-1/2} + C_b(m) \cdot \left(H_y|_{i, j+1/2}^n - \right. \quad (3.41a)$$

$$\left. H_y|_{i-1, j+1/2}^n + H_x|_{i-1/2, j}^n - H_x|_{i-1/2, j+1}^n - J_{\text{source}_z}|_{i-1/2, j+1/2}^n \Delta \right)$$

$$m = \text{MEDIA}_{H_x} \Big|_{i-1/2, j+1}$$

$$\begin{aligned} H_x|_{i-1/2, j+1}^{n+1} &= D_a(m) H_x|_{i-1/2, j+1}^n \\ &+ D_b(m) \cdot \left(E_z|_{i-1/2, j+1/2}^{n+1/2} - E_z|_{i-1/2, j+3/2}^{n+1/2} - M_{\text{source}_x}|_{i-1/2, j+1}^{n+1/2} \Delta \right) \end{aligned} \quad (3.41b)$$

$$m = \text{MEDIA}_{H_y} \Big|_{i, j+1/2}$$

$$\begin{aligned} H_y|_{i, j+1/2}^{n+1} &= D_a(m) H_y|_{i, j+1/2}^n \\ &+ D_b(m) \cdot \left(E_z|_{i+1/2, j+1/2}^{n+1/2} - E_z|_{i-1/2, j+1/2}^{n+1/2} - M_{\text{source}_y}|_{i, j+1/2}^{n+1/2} \Delta \right) \end{aligned} \quad (3.41c)$$

TE_z Mode, Corresponding to the System of (3.14)

Write (3.35a), (3.35b), and (3.36c) without specifying the k index, and then set all partial derivatives of the fields with respect to z equal to zero:

$$m = \text{MEDIA}_{E_x} \Big|_{i, j+1/2}$$

$$\begin{aligned} E_x|_{i, j+1/2}^{n+1/2} &= C_a(m) E_x|_{i, j+1/2}^{n-1/2} \\ &+ C_b(m) \cdot \left(H_z|_{i, j+1, k+1/2}^n - H_z|_{i, j, k+1/2}^n - J_{\text{source}_x}|_{i, j+1/2}^n \Delta \right) \end{aligned} \quad (3.42a)$$

$$m = \text{MEDIA}_{E_y} \Big|_{i-1/2, j+1}$$

$$\begin{aligned} E_y|_{i-1/2, j+1}^{n+1/2} &= C_a(m) E_y|_{i-1/2, j+1}^{n-1/2} \\ &+ C_b(m) \cdot \left(H_z|_{i-1, j+1, k+1/2}^n - H_z|_{i, j+1, k+1/2}^n - J_{\text{source}_y}|_{i-1/2, j+1}^n \Delta \right) \end{aligned} \quad (3.42b)$$

$$m = \text{MEDIA}_{H_z}|_{i,j+1}$$

$$\begin{aligned} H_z|_{i,j+1}^{n+1} &= D_a(m) H_z|_{i,j+1}^n + D_b(m) \cdot \left(E_x|_{i,j+3/2}^{n+1/2} - \right. \\ &\quad \left. E_x|_{i,j+1/2}^{n+1/2} + E_y|_{i-1/2,j+1}^{n+1/2} - E_y|_{i+1/2,j+1}^{n+1/2} - M_{\text{source}_z}|_{i,j+1}^{n+1/2} \Delta \right) \end{aligned} \quad (3.42c)$$

3.6.8 Interpretation as Faraday's and Ampere's Laws in Integral Form

The Yee algorithm for FDTD was originally interpreted as a direct approximation of the pointwise derivatives of Maxwell's time-dependent curl equations by numerical central differences. Although this interpretation is useful for understanding how FDTD models wave propagation away from material interfaces, it sheds little light on what algorithm modifications are needed to properly model the physics of fine geometrical features such as wires, slots, and curved surfaces requiring subcell spatial resolution.

The literature indicates that FDTD modeling can be extended to such features by departing from Yee's original pointwise derivative thinking [2, 3]. As shown in Fig. 3.3, the idea involves starting with a more macroscopic (but still local) combined-field description based upon Ampere's Law and Faraday's Law in *integral* form, implemented on an array of electrically small, spatially orthogonal contours. These contours mesh (intersect) in the manner of links in a chain, providing a geometrical interpretation of the coupling of these two laws. This meshing results in the filling of the FDTD modeled space by a three-dimensional "chain-link" array of intersecting orthogonal contours. The presence of wires, slots, and curved surfaces can be modeled by incorporating appropriate field behavior into the contour and surface integrals used to implement Ampere's and Faraday's Laws at selected meshes, and by deforming contour paths as required to conform with surface curvature.

This approach is intuitively satisfying to an electrical engineer since it permits the FDTD numerical model to deal with physical quantities such as:

- Electromotive forces (EMFs) and magnetomotive forces (MMFs) developed when completing one circuit about a Faraday's or Ampere's Law contour path;
- Magnetic flux and electric displacement current when performing the surface integrations on the patches bounded by the respective contours.

In this section, we demonstrate only the equivalence of the Yee and contour-path interpretations for the free-space case, leaving the modeling of complex spatial features to Chapter 10. For simplicity, FDTD time-stepping expressions are developed for only one field component in Fig. 3.3(a) and one in Fig. 3.3(b). Extension to all the rest is straightforward.

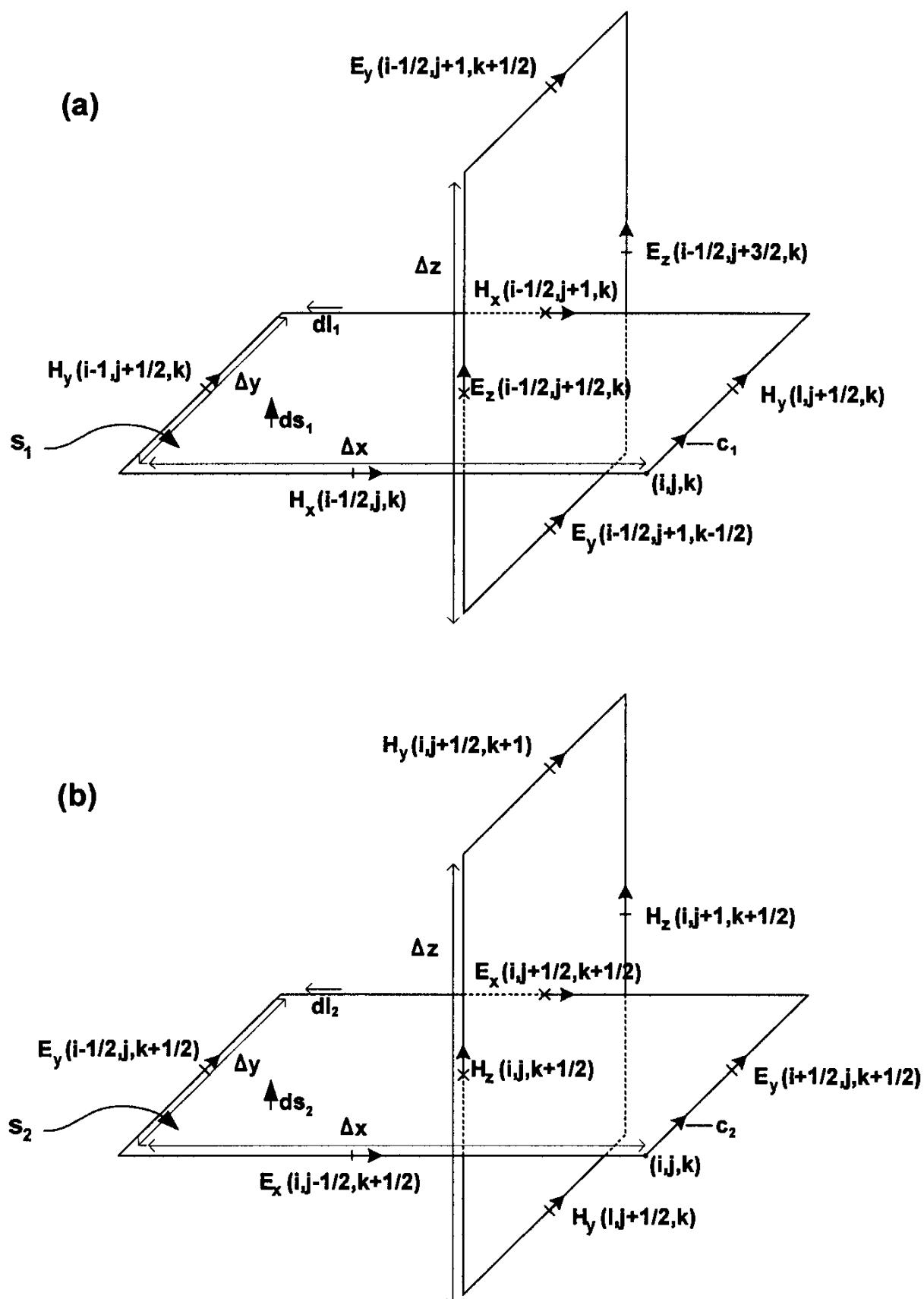


Fig. 3.3 Examples of chain-linked orthogonal contours in the free-space Yee mesh. (a) Ampere's Law for time-stepping E_z ; (b) Faraday's Law for time-stepping H_z . Adapted from: Taflove et al., IEEE Trans. Antennas and Propagation, 1988, pp. 247–257, © 1988 IEEE.

For simplicity, we assume lossless free space with no electric or magnetic current sources. Applying Ampere's Law along contour C_1 in Fig. 3.3(a), and assuming that the field value at a midpoint of one side of the contour equals the average value of that field component along that side, we obtain

$$\frac{\partial}{\partial t} \int_{S_1} \vec{D} \cdot d\vec{S}_1 = \oint_{C_1} \vec{H} \cdot d\vec{\ell}_1 \quad (3.43a)$$

$$\begin{aligned} \frac{\partial}{\partial t} \int_{S_1} \epsilon_0 E_z|_{i-1/2, j+1/2, k} dS_1 &\equiv H_x|_{i-1/2, j, k} \Delta x + H_y|_{i, j+1/2, k} \Delta y \\ &- H_x|_{i-1/2, j+1, k} \Delta x - H_y|_{i-1, j+1/2, k} \Delta y \end{aligned} \quad (3.43b)$$

Now further assume that $E_z|_{i-1/2, j+1/2, k}$ equals the average value of E_z over the surface patch S_1 and that the time derivative can be numerically realized by using a central-difference expression. Then (3.43b) yields

$$\begin{aligned} \epsilon_0 \Delta x \Delta y \left(\frac{E_z|_{i-1/2, j+1/2, k}^{n+1/2} - E_z|_{i-1/2, j+1/2, k}^{n-1/2}}{\Delta t} \right) &= (H_x|_{i-1/2, j, k}^n - H_x|_{i-1/2, j+1, k}^n) \Delta x \\ &+ (H_y|_{i, j+1/2, k}^n - H_y|_{i-1, j+1/2, k}^n) \Delta y \end{aligned} \quad (3.43c)$$

Multiplying both sides by $\Delta t / (\epsilon_0 \Delta x \Delta y)$ and solving for $E_z|_{i-1/2, j+1/2, k}^{n+1/2}$ provides the following time-stepping relation:

$$\begin{aligned} E_z|_{i-1/2, j+1/2, k}^{n+1/2} &= E_z|_{i-1/2, j+1/2, k}^{n-1/2} + (H_x|_{i-1/2, j, k}^n - H_x|_{i-1/2, j+1, k}^n) \Delta t / (\epsilon_0 \Delta y) \\ &+ (H_y|_{i, j+1/2, k}^n - H_y|_{i-1, j+1/2, k}^n) \Delta t / (\epsilon_0 \Delta x) \end{aligned} \quad (3.44)$$

Equation (3.44) is simply the free-space version of (3.29c), the Yee time-stepping equation for E_z that was obtained directly from implementing the curl \vec{H} equation with finite differences. The only difference between (3.29c) and (3.44) is that (3.44) is evaluated at the E_z location $(i-1/2, j+1/2, k)$ whereas (3.29c) is evaluated at the E_z location $(i-1/2, j+1/2, k+1)$ shown in Fig. 3.1.

In an analogous manner, we can apply Faraday's Law along contour C_2 in Fig. 3.3(b) to obtain

$$\frac{\partial}{\partial t} \int_{S_2} \vec{B} \cdot d\vec{S}_2 = - \oint_{C_2} \vec{E} \cdot d\vec{l}_2 \quad (3.45a)$$

$$\begin{aligned} \frac{\partial}{\partial t} \int_{S_2} \mu_0 H_z|_{i,j,k+1/2} dS_2 &\cong -E_x|_{i,j-1/2,k+1/2} \Delta x - E_y|_{i+1/2,j,k+1/2} \Delta y \\ &+ E_x|_{i,j+1/2,k+1/2} \Delta x + E_y|_{i-1/2,j,k+1/2} \Delta y \end{aligned} \quad (3.45b)$$

$$\begin{aligned} \mu_0 \Delta x \Delta y \left(\frac{H_z|_{i,j,k+1/2}^{n+1} - H_z|_{i,j,k+1/2}^n}{\Delta t} \right) &= \left(E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j-1/2,k+1/2}^{n+1/2} \right) \Delta x \\ &+ \left(E_y|_{i-1/2,j,k+1/2}^{n+1/2} - E_y|_{i+1/2,j,k+1/2}^{n+1/2} \right) \Delta y \end{aligned} \quad (3.45c)$$

Multiplying both sides by $\Delta t / (\mu_0 \Delta x \Delta y)$ and solving for $H_z|_{i,j,k+1/2}^{n+1/2}$ provides the following time-stepping relation:

$$\begin{aligned} H_z|_{i,j,k+1/2}^{n+1} &= H_z|_{i,j,k+1/2}^n + \left(E_x|_{i,j+1/2,k+1/2}^{n+1/2} - E_x|_{i,j-1/2,k+1/2}^{n+1/2} \right) \Delta t / (\mu_0 \Delta y) \\ &+ \left(E_y|_{i-1/2,j,k+1/2}^{n+1/2} - E_y|_{i+1/2,j,k+1/2}^{n+1/2} \right) \Delta t / (\mu_0 \Delta x) \end{aligned} \quad (3.46)$$

Equation (3.46) is simply the free-space version of (3.30c), the Yee time-stepping expression for H_z that was obtained directly from implementing the curl \vec{E} equation with finite differences. The only difference between (3.30c) and (3.46) is that (3.46) is evaluated at the H_z location $(i, j, k+1/2)$ whereas (3.30c) is evaluated at the H_z location $(i, j+1, k+1/2)$ shown in Fig. 3.1.

3.6.9 Divergence-Free Nature

As stated earlier, it is crucial for any grid-based solution of Maxwell's curl equations to implicitly enforce (3.3) and (3.4), the Gauss' Law relations for the electric and magnetic fields, which require the absence of free electric and magnetic charge in the source-free space being modeled. We now demonstrate that the Yee space lattice and algorithm satisfy (3.3) for each cell in the lattice, and therefore the lattice as a whole. The proof of the satisfaction of (3.4) by the Yee grid and algorithm is by analogy and is left as an exercise for the student.

We assume lossless free space with no electric or magnetic current sources. (The presence of such current sources is considered in depth in Chapter 5, Section 5.4.) Consider forming the time derivative of the total electric flux over the surface of a single Yee cell of Fig. 3.1:

$$\begin{aligned} \frac{\partial}{\partial t} \iint_{\text{Yee cell}} \bar{D} \cdot d\bar{S} &= \underbrace{\varepsilon_0 \frac{\partial}{\partial t} (E_x|_{i,j+1/2,k+1/2} - E_x|_{i-1,j+1/2,k+1/2})}_{\text{Term 1}} \Delta y \Delta z \\ &+ \underbrace{\varepsilon_0 \frac{\partial}{\partial t} (E_y|_{i-1/2,j+1,k+1/2} - E_y|_{i-1/2,j,k+1/2})}_{\text{Term 2}} \Delta x \Delta z \\ &+ \underbrace{\varepsilon_0 \frac{\partial}{\partial t} (E_z|_{i-1/2,j+1/2,k+1} - E_z|_{i-1/2,j+1/2,k})}_{\text{Term 3}} \Delta x \Delta y \end{aligned} \quad (3.47)$$

Using the Yee algorithm time-stepping relations for the E -field components according to (3.29), we substitute appropriate H -field spatial finite differences for the E -field time derivatives in each term:

$$\begin{aligned} \text{Term 1} &= \left(\frac{H_z|_{i,j+1,k+1/2} - H_z|_{i,j,k+1/2}}{\Delta y} - \frac{H_y|_{i,j+1/2,k+1} - H_y|_{i,j+1/2,k}}{\Delta z} \right) \\ &- \left(\frac{H_z|_{i-1,j+1,k+1/2} - H_z|_{i-1,j,k+1/2}}{\Delta y} - \frac{H_y|_{i-1,j+1/2,k+1} - H_y|_{i-1,j+1/2,k}}{\Delta z} \right) \end{aligned} \quad (3.48a)$$

$$\begin{aligned} \text{Term 2} &= \left(\frac{H_x|_{i-1/2,j+1,k+1} - H_x|_{i-1/2,j+1,k}}{\Delta z} - \frac{H_z|_{i,j+1,k+1/2} - H_z|_{i-1,j+1,k+1/2}}{\Delta x} \right) \\ &- \left(\frac{H_x|_{i-1/2,j,k+1} - H_x|_{i-1/2,j,k}}{\Delta z} - \frac{H_z|_{i,j,k+1/2} - H_z|_{i-1,j,k+1/2}}{\Delta x} \right) \end{aligned} \quad (3.48b)$$

$$\begin{aligned} \text{Term 3} &= \left(\frac{H_y|_{i,j+1/2,k+1} - H_y|_{i-1,j+1/2,k+1}}{\Delta x} - \frac{H_x|_{i-1/2,j+1,k+1} - H_x|_{i-1/2,j,k+1}}{\Delta y} \right) \\ &\quad - \left(\frac{H_y|_{i,j+1/2,k} - H_y|_{i-1,j+1/2,k}}{\Delta x} - \frac{H_x|_{i-1/2,j+1,k} - H_x|_{i-1/2,j,k}}{\Delta y} \right) \end{aligned} \quad (3.48c)$$

For all time steps, this results in

$$\begin{aligned} \frac{\partial}{\partial t} \iint_{\text{Yee cell}} \bar{D} \cdot d\bar{S} &= (\text{Term 1}) \Delta y \Delta z + (\text{Term 2}) \Delta x \Delta z + (\text{Term 3}) \Delta x \Delta y \\ &= 0 \end{aligned} \quad (3.49)$$

Assuming zero initial conditions, the constant zero value of the time derivative of the net electric flux leaving the Yee cell means that this flux never departs from zero:

$$\iint_{\text{Yee cell}} \bar{D}(t) \cdot d\bar{S} = \iint_{\text{Yee cell}} \bar{D}(t=0) \cdot d\bar{S} = 0 \quad (3.50)$$

Therefore, the Yee cell satisfies Gauss' Law for the E -field in charge-free space; hence the Yee algorithm is divergence-free with respect to its E -field computations.

3.7 ALTERNATIVE FINITE-DIFFERENCE GRIDS

Thus far, this chapter has considered several fundamental aspects of the uniform Cartesian Yee space lattice for Maxwell's equations. Since 1966, this lattice and its associated staggered leapfrog time-stepping algorithm have proven to be very flexible, accurate, and robust for a wide variety of engineering problems. However, Yee's staggered, uncollocated arrangement of electromagnetic field components is but one possible alternative in a Cartesian coordinate system [4]. In turn, a Cartesian grid is but one possible arrangement of field components in two and three dimensions. Other possibilities include hexagonal grids in two dimensions and tetrahedron / dual-tetrahedron meshes in three dimensions [4].

It is important to develop criteria for the use of a particular space lattice and time-stepping algorithm to allow optimum selection for a given problem. A key consideration is the capability of rendering the geometry of the structure of interest within the space lattice with sufficient accuracy and detail to obtain meaningful results. A second fundamental consideration is the accuracy by which the algorithm simulates the propagation of electromagnetic waves as they interact with the structure.

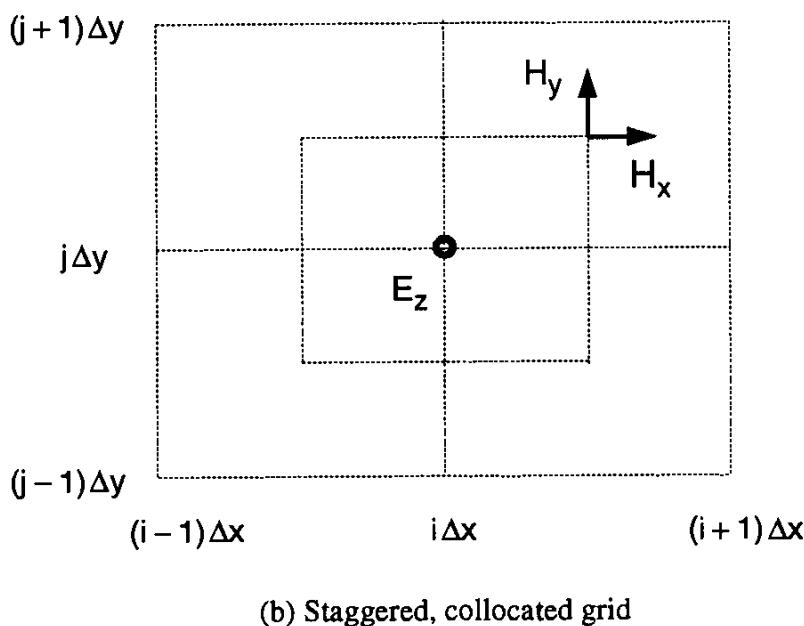
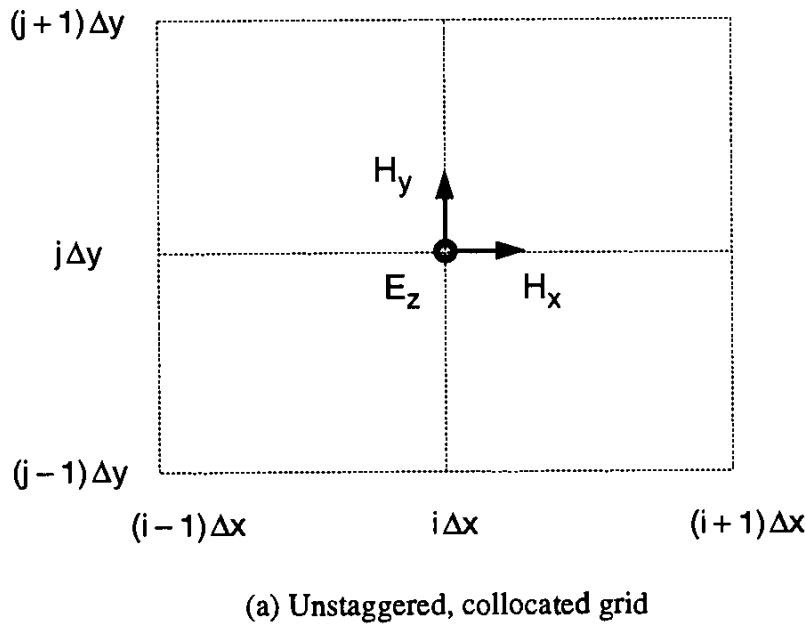


Fig. 3.4 Two Cartesian grids that are alternatives to Yee's arrangement (illustrated in two dimensions for the TM_z case). *Source:* Y. Liu, *J. Computational Physics*, 1996, pp. 396–416.

3.7.1 Cartesian Grids

Fig. 3.4 illustrates two Cartesian grids that are alternatives to Yee's arrangement in two dimensions for the TM_z case [4]: (a) the unstaggered, collocated grid, in which all E and H components are collocated at a single set of grid-cell vertices; and (b) the staggered, collocated grid, in which all E components are collocated at a distinct set of grid-cell vertices that are spatially interleaved with a second distinct set of vertices where all H components are collocated.

Upon applying second-order-accurate central space differences to the TM_z mode equations of (3.13) for the unstaggered, collocated grid of Fig. 3.4(a) (with a lossless material background assumed for simplicity), we obtain [4]:

$$\frac{\partial H_x|_{i,j}}{\partial t} = -\frac{1}{\mu_{i,j}} \cdot \left(\frac{E_z|_{i,j+1} - E_z|_{i,j-1}}{2\Delta y} \right) \quad (3.51a)$$

$$\frac{\partial H_y|_{i,j}}{\partial t} = \frac{1}{\mu_{i,j}} \cdot \left(\frac{E_z|_{i+1,j} - E_z|_{i-1,j}}{2\Delta x} \right) \quad (3.51b)$$

$$\frac{\partial E_z|_{i,j}}{\partial t} = \frac{1}{\epsilon_{i,j}} \cdot \left(\frac{H_y|_{i+1,j} - H_y|_{i-1,j}}{2\Delta x} - \frac{H_x|_{i,j+1} - H_x|_{i,j-1}}{2\Delta y} \right) \quad (3.51c)$$

Similarly, applying second-order-accurate central space differences to the TM_z mode equations of (3.13) for the staggered, collocated grid of Fig. 3.4(b) yields:

$$\frac{\partial H_x|_{i+1/2,j+1/2}}{\partial t} = -\frac{0.5}{\mu_{i+1/2,j+1/2}} \cdot \left[\frac{(E_z|_{i,j+1} + E_z|_{i+1,j+1}) - (E_z|_{i,j} + E_z|_{i+1,j})}{\Delta y} \right] \quad (3.52a)$$

$$\frac{\partial H_y|_{i+1/2,j+1/2}}{\partial t} = \frac{0.5}{\mu_{i+1/2,j+1/2}} \cdot \left[\frac{(E_z|_{i+1,j} + E_z|_{i+1,j+1}) - (E_z|_{i,j} + E_z|_{i,j+1})}{\Delta x} \right] \quad (3.52b)$$

$$\frac{\partial E_z|_{i,j}}{\partial t} = \frac{0.5}{\epsilon_{i,j}} \cdot \left[\begin{aligned} & \frac{(H_y|_{i+1/2,j-1/2} + H_y|_{i+1/2,j+1/2}) - (H_y|_{i-1/2,j-1/2} + H_y|_{i-1/2,j+1/2})}{\Delta x} - \\ & \frac{(H_x|_{i-1/2,j+1/2} + H_x|_{i+1/2,j+1/2}) - (H_x|_{i-1/2,j-1/2} + H_x|_{i+1/2,j-1/2})}{\Delta y} \end{aligned} \right] \quad (3.52c)$$

Reference [4] analyzes the Yee grid and the alternative Cartesian grids of Figs. 3.4(a) and 3.4(b) for a key source of error: the numerical phase-velocity anisotropy. This error, to be discussed in detail in Chapter 4, Section 4.5, is a nonphysical variation of the speed of a numerical wave within an empty grid as a function of its propagation direction. In order to limit this error to less than 0.1%, we require a resolution of 58 points per free-space wavelength λ_0 for the grid of Fig. 3.4(a), 41 points per λ_0 for the grid of Fig. 3.4(b), and only 29 points per λ_0 for the Yee grid [4]. Thus, Yee's grid provides more accurate modeling results than the two alternatives of Fig. 3.4.

3.7.2 Hexagonal Grids

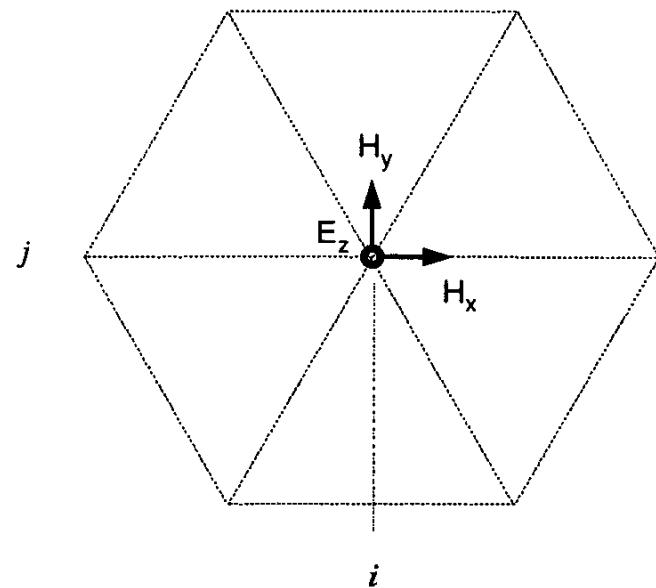
Regular hexagonal grids in two dimensions have been proposed to reduce the numerical phase-velocity anisotropy well below that of Yee's Cartesian mesh [4]. Here, the primary grid is composed of equilateral hexagons having edge length Δs . Each hexagon can be considered to be the union of six equilateral triangles. Connecting the centroids of these triangles yields a second set of regular hexagons that comprises a dual grid.

Fig. 3.5 illustrates for the TM_z case in two dimensions the two principal ways of arranging E and H components in hexagonal grids. Fig. 3.5(a) shows the unstaggered, collocated grid in which Cartesian E_z , H_x , and H_y components are collocated at the vertices of the equilateral triangles. No dual grid is used. Fig. 3.5(b) shows the field arrangement for the staggered, uncollocated grid and its associated dual grid, the latter indicated by the dashed line segments. Here, only E_z components are defined at the vertices of the equilateral triangles, which are the centroids of the hexagonal faces of the dual grid. Magnetic field components H_1 , H_2 , H_3 , etc. are defined to be tangential to, and centered on, the edges of the dual-grid hexagons. These magnetic components are also perpendicular to, and centered on, the edges of the primary-grid triangles. We note that the grid of Fig. 3.5(b) is a direct extension of Yee's interleaved E and H component topology from rectangular to hexagonal cells.

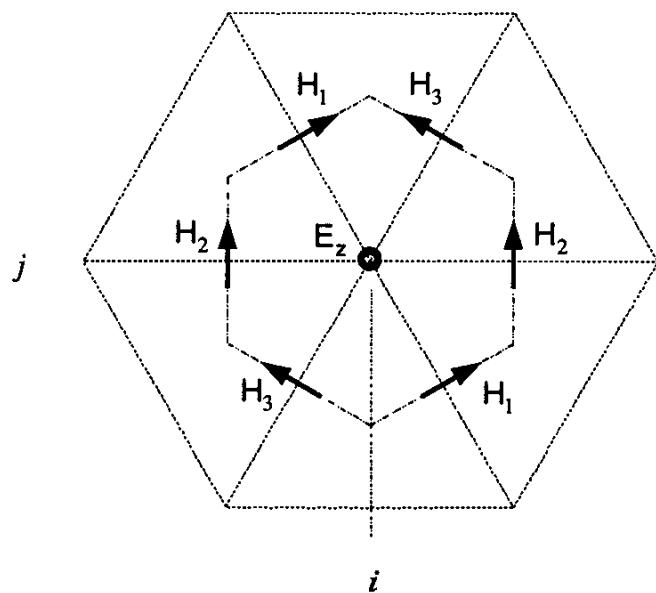
Upon applying second-order-accurate central space differences to the TM_z mode equations of (3.13) for the unstaggered, collocated hexagonal grid of Fig. 3.5(a) (with a lossless material background assumed for simplicity), we obtain [4]:

$$\frac{\partial H_x|_{i,j}}{\partial t} = -\frac{\sqrt{3}}{\mu_{i,j} 6\Delta s} \left(E_z|_{i-1/2,j+0.5\sqrt{3}} + E_z|_{i+1/2,j+0.5\sqrt{3}} - E_z|_{i-1/2,j-0.5\sqrt{3}} - E_z|_{i+1/2,j-0.5\sqrt{3}} \right) \quad (3.53a)$$

$$\frac{\partial H_y|_{i,j}}{\partial t} = \frac{1}{\mu_{i,j} 6\Delta s} \left(2E_z|_{i+1,j} - 2E_z|_{i-1,j} + E_z|_{i+1/2,j+0.5\sqrt{3}} - E_z|_{i-1/2,j+0.5\sqrt{3}} + E_z|_{i+1/2,j-0.5\sqrt{3}} - E_z|_{i-1/2,j-0.5\sqrt{3}} \right) \quad (3.53b)$$



(a) Unstaggered, collocated grid, with no dual grid



(b) Staggered, uncollocated grid and its associated dual grid

Fig. 3.5 Two central-difference hexagonal grids that are alternatives to Yee's arrangement (illustrated in two dimensions for the TM_z case). *Source:* Y. Liu, *J. Computational Physics*, 1996, pp. 396–416.

$$\frac{\partial E_z|_{i,j}}{\partial t} = \frac{1}{\epsilon_{i,j} 6\Delta s} \left(\begin{array}{l} 2H_y|_{i+1,j} - 2H_y|_{i-1,j} + H_y|_{i+1/2,j+0.5\sqrt{3}} \\ - H_y|_{i-1/2,j+0.5\sqrt{3}} + H_y|_{i+1/2,j-0.5\sqrt{3}} - H_y|_{i-1/2,j-0.5\sqrt{3}} \\ - \sqrt{3} H_x|_{i+1/2,j+0.5\sqrt{3}} + \sqrt{3} H_x|_{i+1/2,j-0.5\sqrt{3}} \\ - \sqrt{3} H_x|_{i-1/2,j+0.5\sqrt{3}} + \sqrt{3} H_x|_{i-1/2,j-0.5\sqrt{3}} \end{array} \right) \quad (3.53c)$$

Similarly, applying second-order-accurate central space differences to the TM_z mode equations for the staggered, uncollocated grid of Fig. 3.5(b) yields [4]:

$$\frac{\partial H_1|_{i+1/4,j-0.25\sqrt{3}}}{\partial t} = \frac{1}{\mu_{i+1/4,j-0.25\sqrt{3}} \Delta s} (E_z|_{i+1/2,j-0.5\sqrt{3}} - E_z|_{i,j}) \quad (3.54a)$$

$$\frac{\partial H_2|_{i+1/2,j}}{\partial t} = \frac{1}{\mu_{i+1/2,j} \Delta s} (E_z|_{i+1,j} - E_z|_{i,j}) \quad (3.54b)$$

$$\frac{\partial H_3|_{i+1/4,j+0.25\sqrt{3}}}{\partial t} = \frac{1}{\mu_{i+1/4,j+0.25\sqrt{3}} \Delta s} (E_z|_{i+1/2,j+0.5\sqrt{3}} - E_z|_{i,j}) \quad (3.54c)$$

$$\frac{\partial E_z|_{i,j}}{\partial t} = \frac{2}{\epsilon_{i,j} 3\Delta s} \cdot \left(\begin{array}{l} H_1|_{i+1/4,j-0.25\sqrt{3}} + H_2|_{i+1/2,j} + H_3|_{i+1/4,j+0.25\sqrt{3}} - \\ H_1|_{i-1/4,j+0.25\sqrt{3}} - H_2|_{i-1/2,j} - H_3|_{i-1/4,j-0.25\sqrt{3}} \end{array} \right) \quad (3.54d)$$

We note that the total number of field unknowns for the staggered, uncollocated grid of Fig. 3.5(b) is 33% more than that for the unstaggered grid of Fig. 3.5(a), but the discretization is simpler and the number of total operations is less by about 50%.

Analysis of the numerical dispersion of the hexagonal grids of Figs. 3.5(a) and 3.5(b) shows that the velocity-anisotropy errors are 1/200th and 1/1,200th, respectively, that of the second-order-accurate Yee grid for a grid sampling density of 20 points per free-space wavelength [4]. This represents a large potential advantage in computational accuracy for the hexagonal grids. Additional details are provided in Chapter 4, Section 4.9.3.

3.7.3 Tetradecahedron / Dual-Tetrahedron Mesh in Three Dimensions

In three dimensions, the uniform Cartesian Yee mesh consists of an ordered array of hexahedral unit cells (“bricks”), as shown in Fig. 3.1. This simple arrangement is attractive since the location of every field component in the mesh is easily and compactly specified, and geometry generation can be performed in many cases with paper and pencil.

However, from the discussion of Section 3.7.2, it is clear that constructing a uniform mesh in three dimensions using shapes other than rectangular “bricks” may lead to superior computational accuracy with respect to the reduction of the velocity-anisotropy error. Candidate shapes for unit cells must be capable of assembly in a regular mesh to completely fill space. In addition to the hexahedron, space-filling shapes include the tetradecahedron (truncated octahedron), hexagonal prism, rhombic dodecahedron, and elongated rhombic dodecahedron [4]. We note that the three-dimensional lattice corresponding to the two-dimensional, staggered, uncollocated hexagonal grid of Fig. 3.5(b) is the tetradecahedron / dual-tetrahedron configuration shown in Fig. 3.6. Here, the primary mesh is comprised of tetradecahedral units cells having 6 square faces and 8 regular hexagonal faces. The dual mesh is comprised of tetrahedral cells having isosceles-triangle faces with sides in the ratio of $\sqrt{3}$ to 2.

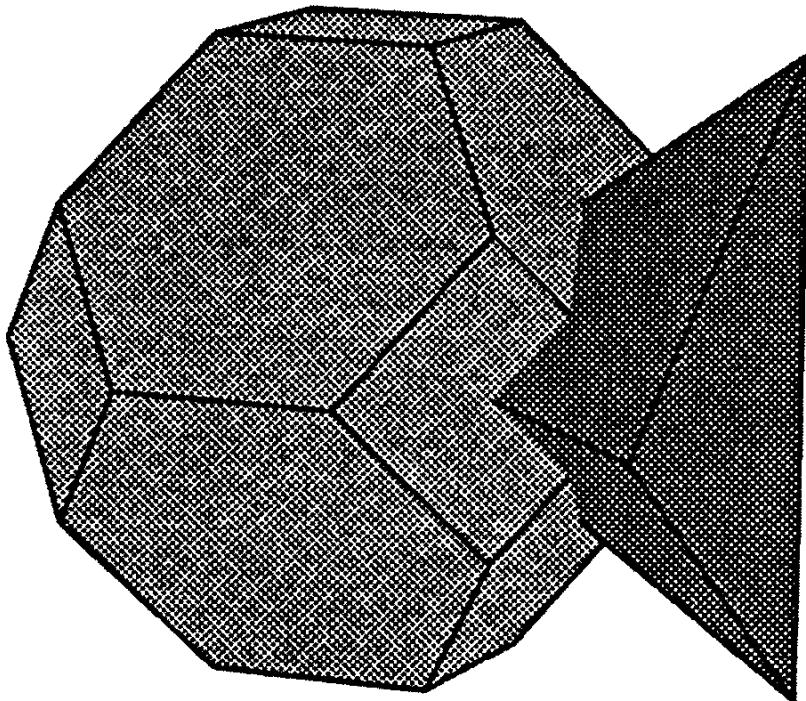


Fig. 3.6 Tetradecahedron and dual-tetrahedron unit cells for the extension of Yee’s method to a regular non-Cartesian mesh in three dimensions. This mesh has very favorable numerical wave-velocity-anisotropy characteristics relative to the Cartesian arrangement of Fig. 3.1. *Source:* Y. Liu, *J. Computational Physics*, 1996, pp. 396–416.

Reference [4] reports a study of the extension of Yee's method to the staggered tetradecahedron/dual-tetrahedron mesh of Fig. 3.6. The algorithm uses a centered finite-difference scheme involving 19 independent unknown field components, wherein 12 are defined on the edges of tetradecahedra, and 7 are defined on the edges of the dual tetrahedra. Similar to the staggered, uncollocated hexagonal grid of Fig. 3.5(b), this mesh has very favorable numerical wave-velocity-anisotropy characteristics relative to its Yee Cartesian counterpart, shown in Fig. 3.1.

Despite this significant accuracy advantage, the usage of the tetradecahedron / dual-tetrahedron mesh by the FDTD community has been quite limited. This is due to the additional complexity in its mesh generation relative to Yee's original Cartesian space lattice. Increased usage of the tetradecahedron / dual-tetrahedron mesh is expected as suitable computer-based mesh-generation tools become more commonly available to the FDTD community.

3.8 SUMMARY

This chapter reviewed Maxwell's equations and the basics of their solution using the Yee algorithm, which forms the foundation of the FDTD method. Specific topics included:

- Review of Maxwell's equations in three dimensions;
- Reduction of Maxwell's equations to two-dimensional TM and TE modes;
- Further reduction of Maxwell's equations to one-dimensional TEM modes;
- Equivalence of Maxwell's equations to the wave equation;
- The Yee algorithm, including:
 - Basic ideas;
 - Finite differences and notation;
 - Finite-difference expressions for Maxwell's equations;
 - Application to space regions with either a continuous variation of material properties or a finite number of distinct materials;
 - Application to nonpermeable materials;
 - Reduction to the two-dimensional TM and TE modes;
 - Interpretation as Faraday's and Ampere's Laws in integral form;
 - Demonstration of its divergence-free nature;
- Alternative Cartesian and hexagonal finite-difference grids.

Since these topics underlie virtually all of the FDTD theory and applications discussed in the remaining chapters of this book, the reader is advised to develop a thorough understanding of each topic.

REFERENCES

- [1] Yee, K. S., "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas and Propagation*, Vol. 14, 1966, pp. 302–307.
- [2] Taflove, A., K. R. Umashankar, B. Beker, F. A. Harfoush, and K. S. Yee, "Detailed FDTD analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens," *IEEE Trans. Antennas and Propagation*, Vol. 36, 1988, pp. 247–257.
- [3] Jurgens, T. G., A. Taflove, K. R. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas and Propagation*, Vol. 40, 1992, pp. 357–366.
- [4] Liu, Y., "Fourier analysis of numerical algorithms for the Maxwell's equations," *J. Computational Physics*, Vol. 124, 1996, pp. 396–416.

PROBLEMS

- 3.1 Show analytically that Gauss' Law in (3.3) and (3.4) for the electric and magnetic fields can be derived from Faraday's Law in (3.1) and Ampere's Law in (3.2) for the case of source-free space.
- 3.2 Show analytically that a scalar wave equation equivalent to (3.20c) can be obtained for the x -directed, y -polarized TEM mode of (3.18).
- 3.3 Construct a computer program that models one-dimensional, x -directed plane-wave propagation in a uniform Yee grid. Use (3.41a) and (3.41c) as your time-stepping algorithm. Assume that $H_x = 0$ and the $j + 1/2$ subscripts can be neglected. Assume that free space is everywhere in the grid, and use the time-step $\Delta t = \Delta x/c$. Terminate the grid in E_z components at its far-left and far-right outer boundaries. Set E_z at the far-left grid boundary to a specific time function such as a unit step, a Gaussian pulse, or a sinusoid. This will radiate a rightward-propagating step, Gaussian, or sinusoidal wave in the grid by the action of the Yee algorithm. Set $E_z = 0$ at the far-right grid boundary to simulate the presence of a *perfect electric conductor* (PEC). Perform visualizations of the E_z and H_y distributions within the grid at a number of time snapshots before and after the propagating wave reaches the far-right grid boundary. Show that the PEC acts as a mirror that reflects the incident wave. Compare the reflection properties of the E_z and H_y components of the wave.
- 3.4 Repeat Problem 3.3 but terminate the grid in an $H_y = 0$ component at its far-right outer boundary to simulate the presence of a *perfect magnetic conductor* (PMC). Compare the wave-reflection properties of the E_z and H_y components of the wave due to the presence of the PMC.
- 3.5 Repeat Problem 3.3 but use the time-step $\Delta t = 0.99\Delta x/c$. Compare your results.

- 3.6 Repeat Problem 3.3 but use the time-step $\Delta t = 1.01\Delta x/c$. Compare your results.
- 3.7 Construct a computer program that models two-dimensional TM_z cylindrical-wave propagation in a uniform Yee grid. Use (3.41a), (3.41b), and (3.41c) as your time-stepping algorithm. Assume square unit cells $\Delta x = \Delta y \equiv \Delta$, free space everywhere in the grid, and a time-step $\Delta t = \Delta / (c\sqrt{2})$. Terminate the grid in $E_z = 0$ components at its outer boundaries, thereby simulating the presence of PECs there. Set a single E_z component located at the center of the grid to a specific time function such as a unit step, a Gaussian pulse, or a sinusoid. This will generate a radially-propagating step, Gaussian, or sinusoidal wave in the grid by the action of the Yee algorithm. Perform visualizations of the E_z , H_x , and H_y fields of the outgoing wave distributed within the grid at a number of time snapshots before and after the wave reaches the outer grid boundary. Compare the spatial-symmetry properties of these fields with respect to the center of the grid where the excitation is applied.
- 3.8 Repeat Problem 3.7, but fill the grid with a uniform, electrically conductive medium. Vary the conductivity upward from zero and observe how the outgoing wave is attenuated.
- 3.9 Repeat Problem 3.7 but now use the time-step $\Delta t = 1.0005\Delta / (c\sqrt{2})$.
- 3.10 Repeat Problem 3.7 for the two-dimensional TE_z mode. Use (3.42a), (3.42b), and (3.42c) as your time-stepping algorithm. Terminate the grid in $E_x = 0$ and $E_y = 0$ components at its outer boundaries, thereby simulating the presence of PECs there. Set a single H_z component located at the center of the grid to a specific time function such as a unit step, a Gaussian pulse, or a sinusoid. Perform visualizations of the H_z , E_x , and E_y fields of the outgoing wave distributed within the grid at a number of time snapshots before and after the wave reaches the outer grid boundary. Compare the spatial-symmetry properties of these fields with respect to the center of the grid.
- 3.11 Repeat Problem 3.10, but now use PMC outer grid boundaries ($H_z = 0$).
- 3.12 Verify the divergence-free nature of the Yee grid and algorithm for its computed magnetic fields.

Chapter 4

Numerical Dispersion and Stability

4.1 INTRODUCTION

The FDTD algorithms for Maxwell's curl equations reviewed in Chapter 3 cause nonphysical dispersion of the simulated waves in a free-space computational lattice. That is, the phase velocity of numerical wave modes can differ from c by an amount varying with the wavelength, direction of propagation in the grid, and grid discretization. An intuitive way to view this phenomenon is that the FDTD algorithm embeds the electromagnetic wave interaction structure of interest in a tenuous "numerical aether" having properties very close to vacuum, but not quite. This "aether" causes propagating numerical waves to accumulate delay or phase errors that can lead to nonphysical results such as broadening and ringing of pulsed waveforms, imprecise cancellation of multiple scattered waves, anisotropy, and pseudorefraction. Numerical dispersion is a factor in FDTD modeling that must be accounted for to understand its operation and its accuracy limits, especially for electrically large structures.

In addition, the FDTD algorithms for Maxwell's equations reviewed in Chapter 3 require that the time step Δt has a specific bound relative to the lattice space increments Δx , Δy , and Δz . This bound is necessary to avoid numerical instability, an undesirable possibility with explicit differential equation solvers that can cause the computed results to spuriously increase without limit as time-marching continues.

This chapter derives the key relations for numerical dispersion and stability applicable to FDTD modeling of Maxwell's equations in multiple dimensions. These derivations build upon the analyses introduced in Chapter 2 for the one-dimensional scalar wave equation. Additional results from the literature will be presented for non-Cartesian space lattices and emerging low-dispersion FDTD techniques.

4.2 DERIVATION OF THE NUMERICAL DISPERSION RELATION FOR TWO-DIMENSIONAL WAVE PROPAGATION

We begin our discussion of numerical dispersion with an analysis of the Yee algorithm implementation of (3.15a), (3.15b), and (3.15c), the field equations for the two-dimensional TM_z mode. It can be easily shown that the numerical dispersion relation obtained is valid for any two-dimensional TM or TE mode. Further, our analysis serves as a convenient starting point for extension to three-dimensional problems as well as simplification to one-dimensional problems. Assuming for simplicity no magnetic or electric loss, the system of (3.15) yields

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu} \frac{\partial E_z}{\partial y} \quad (4.1a)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \frac{\partial E_z}{\partial x} \quad (4.1b)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \quad (4.1c)$$

To further simplify the problem, we assume that the FDTD modeling space is filled with homogeneous material having no variation of μ or ϵ with position in the grid. Then the finite-difference expressions for the TM_z case are given by the following:

$$\frac{H_x|_{i,j+1/2}^{n+1/2} - H_x|_{i,j+1/2}^{n-1/2}}{\Delta t} = -\frac{1}{\mu_{i,j+1/2}} \left(\frac{E_z|_{i,j+1}^n - E_z|_{i,j}^n}{\Delta y} \right) \quad (4.2a)$$

$$\frac{H_y|_{i+1/2,j}^{n+1/2} - H_y|_{i+1/2,j}^{n-1/2}}{\Delta t} = \frac{1}{\mu_{i+1/2,j}} \left(\frac{E_z|_{i+1,j}^n - E_z|_{i,j}^n}{\Delta x} \right) \quad (4.2b)$$

$$\frac{E_z|_{i,j}^{n+1} - E_z|_{i,j}^n}{\Delta t} = \frac{1}{\epsilon_{i,j}} \left(\frac{H_y|_{i+1/2,j}^{n+1/2} - H_y|_{i-1/2,j}^{n+1/2}}{\Delta x} - \frac{H_x|_{i,j+1/2}^{n+1/2} - H_x|_{i,j-1/2}^{n+1/2}}{\Delta y} \right) \quad (4.2c)$$

The basic procedure for the numerical dispersion analysis involves substitution of a plane, monochromatic, traveling-wave trial solution into the finite-difference system of (4.2). After algebraic manipulation, an equation is derived that relates the numerical wavevector components, the wave frequency, the time-step, and the grid space

increments. This equation, the numerical dispersion relation, can be solved for a variety of grid discretizations, wavevectors, and wave frequencies to illustrate the key nonphysical modeling results associated with numerical dispersion.

Initiating this procedure, we assume the following plane, monochromatic, traveling-wave trial solution for the TM_z mode:

$$E_z|_{I,J}^n = E_{z_0} e^{j(\omega n \Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)} \quad (4.3a)$$

$$H_x|_{I,J}^n = H_{x_0} e^{j(\omega n \Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)} \quad (4.3b)$$

$$H_y|_{I,J}^n = H_{y_0} e^{j(\omega n \Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y)} \quad (4.3c)$$

where \tilde{k}_x and \tilde{k}_y are the x - and y -components of the numerical wavevector and ω is the wave angular frequency. Substituting the traveling-wave expressions of (4.3) into the finite-difference equations of (4.2) yields, after simplification, the following relations:

$$H_{x_0} = \frac{\Delta t E_{z_0}}{\mu \Delta y} \cdot \frac{\sin(\tilde{k}_y \Delta y / 2)}{\sin(\omega \Delta t / 2)} \quad (4.4a)$$

$$H_{y_0} = -\frac{\Delta t E_{z_0}}{\mu \Delta x} \cdot \frac{\sin(\tilde{k}_x \Delta x / 2)}{\sin(\omega \Delta t / 2)} \quad (4.4b)$$

$$E_{z_0} \sin\left(\frac{\omega \Delta t}{2}\right) = \frac{\Delta t}{\epsilon} \left[\frac{H_{x_0}}{\Delta y} \sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) - \frac{H_{y_0}}{\Delta x} \sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) \right] \quad (4.4c)$$

Upon substituting H_{x_0} of (4.4a) and H_{y_0} of (4.4b) into (4.4c), we obtain

$$\left[\frac{1}{c \Delta t} \sin\left(\frac{\omega \Delta t}{2}\right) \right]^2 = \left[\frac{1}{\Delta x} \sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) \right]^2 + \left[\frac{1}{\Delta y} \sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) \right]^2 \quad (4.5)$$

where $c = 1/\sqrt{\mu \epsilon}$ is the speed of light in the material being modeled. Equation (4.5) is the general numerical dispersion relation of the Yee algorithm for the TM_z mode.

We shall consider the important special case of a square-cell grid having $\Delta x = \Delta y = \Delta$. Then, we can use the Chapter 2 definitions of the Courant stability factor $S = c \Delta t / \Delta$ and the grid sampling density $N_\lambda = \lambda_0 / \Delta$ to rewrite (4.5) in a more useful form:

$$\frac{1}{S^2} \sin^2\left(\frac{\pi S}{N_\lambda}\right) = \sin^2\left(\frac{\Delta \cdot \tilde{k} \cos \phi}{2}\right) + \sin^2\left(\frac{\Delta \cdot \tilde{k} \sin \phi}{2}\right) \quad (4.6)$$

where ϕ is the propagation direction of the numerical wave with respect to the grid's x -axis. To obtain the numerical dispersion relation for the one-dimensional wave-propagation case, we can assume without loss of generality that $\phi = 0$ in (4.6). Then, (4.6) reduces to

$$\frac{1}{S} \sin\left(\frac{\pi S}{N_\lambda}\right) = \sin\left(\frac{\tilde{k} \Delta}{2}\right) \quad (4.7a)$$

or equivalently

$$\tilde{k} = \frac{2}{\Delta} \sin^{-1}\left[\frac{1}{S} \sin\left(\frac{\pi S}{N_\lambda}\right)\right] \quad (4.7b)$$

4.3 EXTENSION TO THREE DIMENSIONS

The dispersion analysis presented above is now extended to the full three-dimensional case involving all six coupled E - and H -field vector components. For convenience, this section utilizes the compact vector notation for Maxwell's equations introduced in [1].

Following [1], we consider a normalized, lossless region of space with $\mu = 1$, $\epsilon = 1$, $\sigma = 0$, $\sigma^* = 0$, and $c = 1$. Letting $j = \sqrt{-1}$, we rewrite Maxwell's equations in compact form as

$$j \nabla \times (\vec{H} + j\vec{E}) = \frac{\partial}{\partial t} (\vec{H} + j\vec{E}) \quad (4.8a)$$

or more simply as

$$j \nabla \times \vec{V} = \frac{\partial \vec{V}}{\partial t} \quad (4.8b)$$

where $\vec{V} = \vec{H} + j\vec{E}$. Substituting the vector-field traveling-wave expression

$$\vec{V}|_{I,J,K}^n = \vec{V}_0 e^{j(\omega n \Delta t - \tilde{k}_x I \Delta x - \tilde{k}_y J \Delta y - \tilde{k}_z K \Delta z)} \quad (4.9)$$

into the Yee space-time central-differencing realization of (4.8b), we obtain

$$\begin{aligned}
& \left[\frac{\hat{x}}{\Delta x} \sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) + \frac{\hat{y}}{\Delta y} \sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) + \frac{\hat{z}}{\Delta z} \sin\left(\frac{\tilde{k}_z \Delta z}{2}\right) \right] \times \vec{V}|_{I,J,K}^n \\
&= \frac{-j}{\Delta t} \vec{V}|_{I,J,K}^n \sin\left(\frac{\omega \Delta t}{2}\right)
\end{aligned} \tag{4.10}$$

where \hat{x} , \hat{y} , and \hat{z} are unit vectors in the x -, y -, and z -coordinate directions. After performing the vector cross product in (4.10) and writing out the x , y , and z vector component equations, we obtain a homogeneous system (zero right-hand side) of three equations in the unknowns V_x , V_y , and V_z . Setting the determinant of this system equal to zero results in

$$\begin{aligned}
\left[\frac{1}{\Delta t} \sin\left(\frac{\omega \Delta t}{2}\right) \right]^2 &= \left[\frac{1}{\Delta x} \sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) \right]^2 + \left[\frac{1}{\Delta y} \sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) \right]^2 \\
&+ \left[\frac{1}{\Delta z} \sin\left(\frac{\tilde{k}_z \Delta z}{2}\right) \right]^2
\end{aligned} \tag{4.11}$$

Finally, we denormalize to a nonunity c and obtain the general form of the numerical dispersion relation for the full-vector-field Yee algorithm in three dimensions:

$$\begin{aligned}
\left[\frac{1}{c \Delta t} \sin\left(\frac{\omega \Delta t}{2}\right) \right]^2 &= \left[\frac{1}{\Delta x} \sin\left(\frac{\tilde{k}_x \Delta x}{2}\right) \right]^2 + \left[\frac{1}{\Delta y} \sin\left(\frac{\tilde{k}_y \Delta y}{2}\right) \right]^2 \\
&+ \left[\frac{1}{\Delta z} \sin\left(\frac{\tilde{k}_z \Delta z}{2}\right) \right]^2
\end{aligned} \tag{4.12}$$

This equation is seen to reduce to (4.5), the numerical dispersion relation for the two-dimensional TM_z mode, simply by letting $\tilde{k}_z = 0$.

4.4 COMPARISON WITH THE IDEAL DISPERSION CASE

In contrast to the numerical dispersion relation of (4.12), the analytical (ideal) dispersion relation for a physical plane wave propagating in three dimensions in a homogeneous lossless medium is simply

$$\left(\frac{\omega}{c} \right)^2 = (k_x)^2 + (k_y)^2 + (k_z)^2 \tag{4.13}$$

Although at first glance (4.12) bears little resemblance to the ideal case of (4.13), we can easily show that the two dispersion relations are identical in the limit as Δx , Δy , Δz , and Δt approach zero. Qualitatively, this suggests that numerical dispersion can be reduced to any degree that is desired if we only use fine enough FDTD gridding.

It can also be shown that (4.12) reduces to (4.13) if the Courant factor and the direction of wave propagation are suitably chosen. For example, reduction to the ideal dispersion case can be demonstrated for a numerical plane wave propagating along a diagonal of a three-dimensional cubic lattice ($\tilde{k}_x = k_y = \tilde{k}_z = \tilde{k}/\sqrt{3}$) if $S = 1/\sqrt{3}$. Similarly, ideal dispersion results for a numerical plane wave propagating along a diagonal of a two-dimensional square grid ($\tilde{k}_x = k_y = \tilde{k}/\sqrt{2}$) if $S = 1/\sqrt{2}$. Finally, ideal dispersion results for any numerical wave in a uniform one-dimensional grid if $S = 1$, the magic time-step. These reductions to the ideal case have little practical value for two- and three-dimensional simulations, occurring uniquely for diagonal propagation and not some more general set of wave directions. However, the reduction to ideal dispersion in one dimension is very interesting, since it implies that the Yee algorithm (based upon finite-difference approximations) can *exactly* solve the continuous one-dimensional wave equation.

4.5 ANISOTROPY OF THE NUMERICAL PHASE VELOCITY

This section probes a key implication of numerical dispersion relations (4.5) and (4.12). Namely, numerical waves in a two- or three-dimensional Yee space lattice have a propagation velocity that is dependent upon the direction of wave propagation. The space lattice thus represents an anisotropic medium.

Our strategy in developing an understanding of this phenomenon is to first calculate sample values of the numerical phase velocity \tilde{v}_p versus wave-propagation direction ϕ in order to estimate the magnitude of the problem. Then, we will conduct an appropriate analysis to examine the issue more deeply.

4.5.1 Sample Values of Numerical Phase Velocity

For simplicity, we start with the simplest possible situation where numerical phase-velocity anisotropy arises: two-dimensional TM_z modes propagating in a square-cell grid. Dispersion relation (4.6) can be solved directly for \tilde{k} for propagation along the major axes of the grid: $\phi = 0^\circ$, 90° , 180° , and 270° . For this case, the solution for \tilde{k} is given by (4.7b), which is repeated here for convenience:

$$\tilde{k} = \frac{2}{\Delta} \sin^{-1} \left[\frac{1}{S} \sin \left(\frac{\pi S}{N_\lambda} \right) \right] \quad : \text{propagation along major grid axes} \quad (4.14a)$$

The corresponding numerical phase velocity is given by

$$\tilde{v}_p = \frac{\omega}{\tilde{k}} = \frac{\pi}{N_\lambda \sin^{-1} \left[\frac{1}{S} \sin \left(\frac{\pi S}{N_\lambda} \right) \right]} c \quad : \text{propagation along major grid axes} \quad (4.14b)$$

Dispersion relation (4.6) can also be solved directly for \tilde{k} for propagation along the diagonals of the grid: $\phi = 45^\circ, 135^\circ, 225^\circ$, and 315° . For this case, the solution for \tilde{k} and \tilde{v}_p is given by

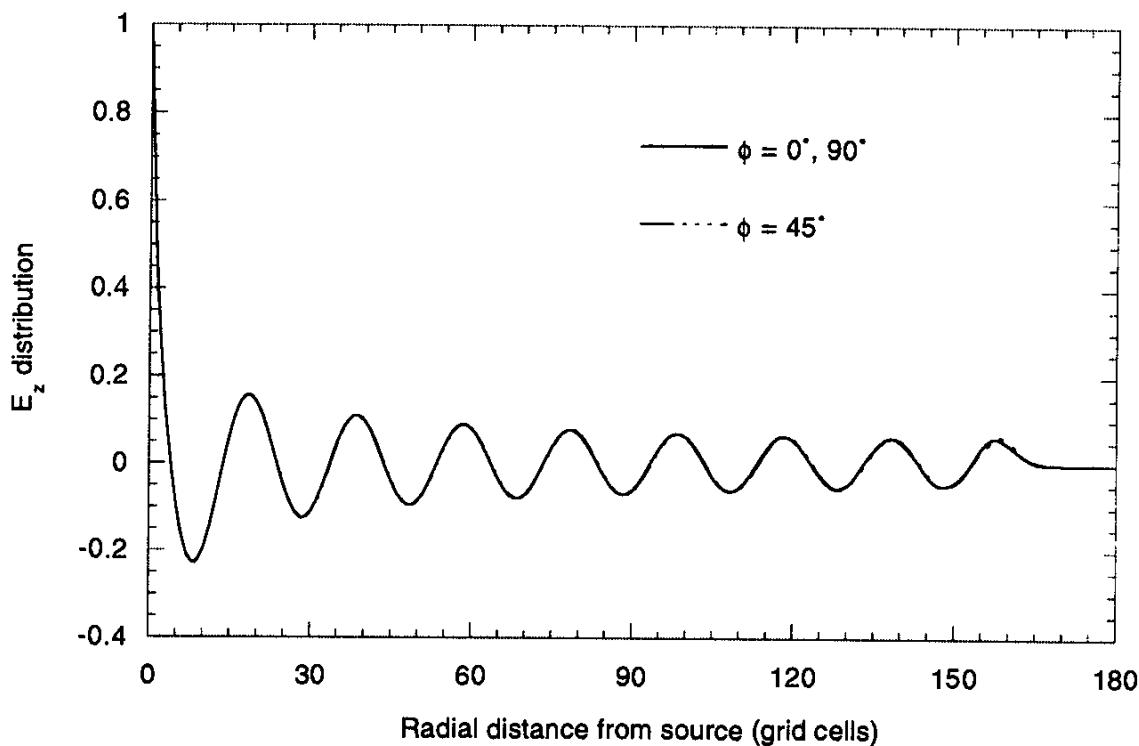
$$\tilde{k} = \frac{2\sqrt{2}}{\Delta} \sin^{-1} \left[\frac{1}{S\sqrt{2}} \sin \left(\frac{\pi S}{N_\lambda} \right) \right] \quad : \text{propagation along grid diagonals} \quad (4.15a)$$

$$\tilde{v}_p = \frac{\pi}{N_\lambda \sqrt{2} \sin^{-1} \left[\frac{1}{S\sqrt{2}} \sin \left(\frac{\pi S}{N_\lambda} \right) \right]} c \quad : \text{propagation along grid diagonals} \quad (4.15b)$$

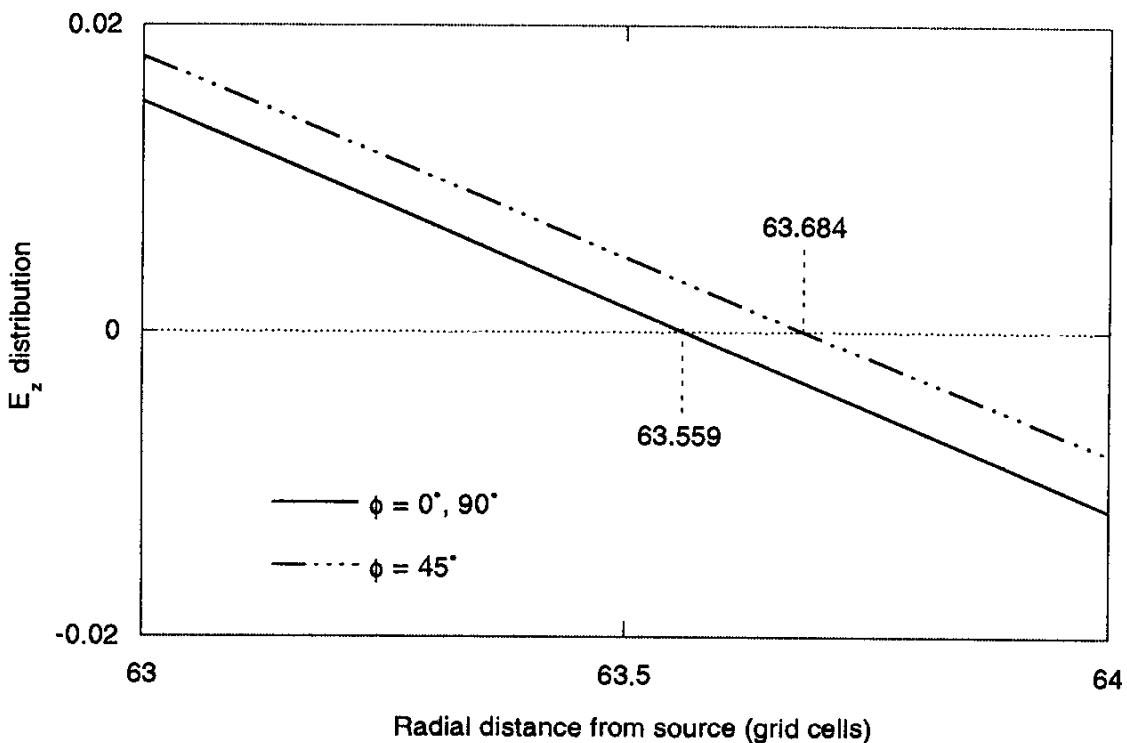
As an example of values arising from these expressions, assume a grid having $S = 0.5$ and $N_\lambda = 20$. Then (4.14b) and (4.15b) yield numerical phase velocities of $0.996892c$ and $0.998968c$, respectively. Both of these velocities are less than c , indicating a phase lag relative to a physical wave propagating in free space at these angles. Further, the velocities are not equal. The implication is that a sinusoidal numerical wave propagating obliquely within this grid should have a speed that is $0.998968 / 0.996892 = 1.00208$ times that of a wave propagating along the major grid axes. This represents a velocity anisotropy of about 0.2% between oblique and along-axis numerical wave propagation.

We now demonstrate that this theoretical anisotropy of the numerical phase velocity actually appears in FDTD simulations. Fig. 4.1 presents FDTD modeling results for a radially outward-propagating sinusoidal cylindrical wave in a two-dimensional TM_z grid. The grid is configured with 360×360 square cells with $\Delta x = \Delta y = \Delta = 1.0$. Starting at time step $n = 1$, a unity-amplitude sinusoidal excitation is provided to a single E_z field component at the center point of the grid. Given that the radiated numerical mode for this case has an easily specified free-space wavelength λ_0 , we choose a grid-sampling density relative to this wavelength as $N_\lambda = 20$. Further, we choose the Courant factor $S = 0.5$. This permits direct comparison of the FDTD modeling results with the theoretical results for anisotropy of \tilde{v}_p , discussed in the paragraph immediately above.

In Fig. 4.1(a), we graph snapshots of the E_z field distribution vs. radial distance from the source at the center of the grid. Here, field observations are made along cuts through the grid passing through the source point and either parallel to the principal grid axes $\phi = 0^\circ, 90^\circ$ or parallel to the grid diagonal $\phi = 45^\circ$. (We note that, by the 90° rotational symmetry of the Cartesian grid geometry, identical field distributions are



(a) Comparison of calculated wave propagation along the grid axes and along a grid diagonal.



(b) Expanded view of (a) at distances between 63 and 64 grid cells from the source.

Fig. 4.1 Effect of numerical dispersion upon a radially propagating cylindrical wave in a 2-D TM Yee grid. The grid is excited at its center point by applying a unity-amplitude sinusoidal time function to a single E_z field component. $S = 0.5$ and the grid sampling density is $N_\lambda = 20$.